```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         January 17, 2004, 01:15:10 ; Search time 10931 Seconds
                                              (without alignments)
                                              5344.337 Million cell updates/sec

Title:          US-10-020-095-4
Perfect score:  7348
Sequence:       1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters:      5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10020095/runat_15012004_163822_18707/app_query.fasta_1.1607
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10020095_@CGN_1_1_7794_@runat_15012004_163822_18707 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

SUMMARY
SEQ ID NO: 3
SEQ ID NO: 4
REV. TRANS. SEQ ID NO: 4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7326.5	99.7	4335	6	AX537610	AX537610 Sequence
2	7326.5	99.7	4761	6	AX534934	AX534934 Sequence
3	7326.5	99.7	5883	9	AF410459	AF410459 Homo sapi
4	7326.5	99.7	5895	6	AX534938	AX534938 Sequence
5	7317.5	99.6	4335	6	AX537612	AX537612 Sequence
6	7317.5	99.6	4761	6	AX534936	AX534936 Sequence
7	7317.5	99.6	5895	6	AX534940	AX534940 Sequence
8	5470	74.4	3535	6	AX534942	AX534942 Sequence
9	5461	74.3	3535	6	AX534944	AX534944 Sequence
10	5360	72.9	4735	10	AY083458	AY083458 Mus muscu
11	5360	72.9	5644	10	BC052443	BC052443 Mus muscu
12	4305.5	58.6	2938	9	HSM805581	AL834478 Homo sapi
13	3360	45.7	2273	9	AK095888	AK095888 Homo sapi
14	1854.5	25.2	4527	6	AX054953	AX054953 Sequence
15	1850	25.2	4560	6	AX054954	AX054954 Sequence
16	1779.5	24.2	5603	3	CIN431688	AJ431688 Ciona int
17	1692	23.0	4578	3	AY122084	AY122084 Drosophil
18	1692	23.0	4579	3	DME269539	AJ269539 Drosophil
19	1638	22.3	4678	3	D83196	D83196 Limulus sp.
20	1542.5	21.0	5817	9	AB033109	AB033109 Homo sapi
21	1542.5	21.0	5852	9	AY101765	AY101765 Homo sapi
22	1491	20.3	4545	10	GPIAMA	D84338 Cavia porce
23	1481.5	20.2	5210	9	HSM803446	AL832139 Homo sapi
24	1463.5	19.9	4598	3	DME269541	AJ269541 Drosophil
25	1460	19.9	4220	3	DME269538	AJ269538 Drosophil
26	1459	19.9	4944	3	AF538967	AF538967 Ornithodo
27	1455	19.8	4636	10	RATA1I3A	J03552 Rat plasma
28	1441	19.6	4577	6	AR241630	AR241630 Sequence
29	1440	19.6	4425	6	AX054943	AX054943 Sequence
30	1440	19.6	4577	6	AX146876	AX146876 Sequence
31	1440	19.6	4577	6	AX409589	AX409589 Sequence
32	1440	19.6	4577	9	HUMA2M	M11313 Human alpha
33	1440	19.6	4617	9	BC040071	BC040071 Homo sapi
34	1440	19.6	4809	6	AX281769	AX281769 Sequence
35	1437	19.6	4801	9	HSM804792	AL833479 Homo sapi
36	1435.5	19.5	4841	5	LARA2M	D13567 Lampetra ja
37	1434	19.5	4569	6	A21185	A21185 H.sapiens m
38	1432.5	19.5	4622	10	AY185125	AY185125 Mus muscu
39	1428.5	19.4	6377	3	AB050668	AB050668 Branchios
40	1425.5	19.4	4650	3	DME269540	AJ269540 Drosophil
41	1425.5	19.4	4913	3	AY069050	AY069050 Drosophil
42	1422	19.4	4488	6	AX497234	AX497234 Sequence
43	1422	19.4	4488	6	AX574598	AX574598 Sequence
44	1421	19.3	4599	6	A21187	A21187 H.sapiens m
45	1420	19.3	4672	10	BC051037	BC051037 Mus muscu

ALIGNMENTS

RESULT 1

```
AX537610
LOCUS       AX537610                4335 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO02070738.
ACCESSION   AX537610
VERSION     AX537610.1  GI:25269531
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Ouwehand,W.
  TITLE     Diagnosis and treatment of blood disorders
  JOURNAL   Patent: WO 02070738-A 1 12-SEP-2002;
            Schuh, Andre (CA)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..>4335
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57496.1"
                     /db_xref="GI:25269532"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1296 a    837 c    922 g   1280 t
ORIGIN

Alignment Scores:
Pred. No.:              0            Length:         4335
Score:                  7326.50      Matches:        1427
Percent Similarity:     98.82%       Conservative:   1
Best Local Similarity:  98.75%       Mismatches:     0
Query Match:            99.71%       Indels:         17
DB:                     6            Gaps:           1

US-10-020-095-4 (1-1428) x AX537610 (1-4335)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy       421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy       441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy       461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy       481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy       501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy       521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy       541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy       561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy       581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy       601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy       621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy       641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy       661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy       681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy       701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy       721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy       741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy       761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy       781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
             ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db      2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400

Qy       801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy       821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy       841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy       861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy       881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy       901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy       921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy       941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy       961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy       981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy      1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy      1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy      1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy      1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy      1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300

Qy      1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy      1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
```

```
Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy       1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                        ||||||||||||||||||
Db       3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4321 ATGGAACTTTGGCTG 4335
```

```
RESULT 2
AX534934
LOCUS       AX534934                4761 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 1 from Patent WO02070696.
ACCESSION   AX534934
VERSION     AX534934.1  GI:25261474
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 1 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1..4761
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113..4450
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57231.1"
                     /db_xref="GI:25261475"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1420 a    912 c   1016 g   1413 t
ORIGIN

Alignment Scores:
Pred. No.:              0             Length:          4761
Score:                  7326.50       Matches:         1427
Percent Similarity:     98.82%        Conservative:    1
Best Local Similarity:  98.75%        Mismatches:      0
Query Match:            99.71%        Indels:          17
DB:                     6             Gaps:            1

US-10-020-095-4 (1-1428) x AX534934 (1-4761)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAAGCCAGTG 892

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              |||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572
```

```
Qy         821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy         841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy         861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy         881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy         901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy         921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy         941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy         961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy         981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy        1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy        1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy        1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy        1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy        1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy        1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy        1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy        1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy        1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy        1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy        1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy        1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                         ||||||||||||||||||
Db        3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy        1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy        1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy        1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy        1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy        1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy        1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy        1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy        1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy        1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy        1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy        1424 MetGluLeuTrpLeu 1428
               |||||||||||||||
Db        4433 ATGGAACTTTGGCTG 4447

RESULT 3
AF410459
LOCUS       AF410459                5883 bp    mRNA    linear   PRI 02-MAR-2002
DEFINITION  Homo sapiens CD109 (CD109) mRNA, complete cds.
ACCESSION   AF410459
VERSION     AF410459.1  GI:19071208
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5883)
  AUTHORS   Lin,M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
            Wu,X.F. and Schuh,A.C.
  TITLE     Cell surface antigen CD109 is a novel member of the alpha(2)
            macroglobulin/C3, C4, C5 family of thioester-containing proteins
  JOURNAL   Blood 99 (5), 1683-1691 (2002)
```

```
MEDLINE     21849742
 PUBMED     11861284
REFERENCE   2  (bases 1 to 5883)
 AUTHORS    Lin,M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
            Wu,X.-F. and Schuh,A.C.
 TITLE      Direct Submission
 JOURNAL    Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's
            College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada
FEATURES             Location/Qualifiers
     source          1. .5883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="K1"
     gene            1. .5883
                     /gene="CD109"
     5'UTR           1. .112
                     /gene="CD109"
     CDS             113. .4450
                     /gene="CD109"
                     /note="associated with the Gov alloantigen system"
                     /codon_start=1
                     /product="CD109"
                     /protein_id="AAL84159.1"
                     /db_xref="GI:19071209"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
     3'UTR           4451.. .5883
                     /gene="CD109"
     misc_feature    4748
                     /gene="CD109"
                     /note="alternative polyA site found on clone K1"
BASE COUNT     1697 a   1173 c   1241 g   1772 t
ORIGIN

Alignment Scores:
Pred. No.:              0             Length:         5883
Score:                  7326.50       Matches:        1427
Percent Similarity:     98.82%        Conservative:   1
Best Local Similarity:  98.75%        Mismatches:     0
Query Match:            99.71%        Indels:         17
DB:                     9             Gaps:           1

US-10-020-095-4 (1-1428) x AF410459 (1-5883)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy         401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy         421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy         441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy         461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy         481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy         501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy         521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy         541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy         561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy         581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy         601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy         621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy         641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy         661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy         681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy         701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy         721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy         741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy         761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy         781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
               ||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||
Db        2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy         801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy         821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy         841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy         861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy         881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy         901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy         921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy         941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy         961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy         981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy        1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy        1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy        1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy        1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy        1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy        1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472
```

```
Qy     1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy     1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy     1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy     1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy     1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy     1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                      ||||||||||||||||||
Db     3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy     1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy     1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy     1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy     1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy     1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy     1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy     1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy     1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy     1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy     1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy     1424 MetGluLeuTrpLeu 1428
            |||||||||||||||
Db     4433 ATGGAACTTTGGCTG 4447
```

```
RESULT 4
AX534938
LOCUS       AX534938                 5895 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 5 from Patent WO02070696.
ACCESSION   AX534938
VERSION     AX534938.1  GI:25261482
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 5 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1..5895
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113..4450
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57233.1"
                     /db_xref="GI:25261483"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1709 a   1173 c   1241 g   1772 t
ORIGIN

Alignment Scores:
Pred. No.:                0               Length:          5895
Score:                    7326.50         Matches:         1427
Percent Similarity:       98.82%          Conservative:    1
Best Local Similarity:    98.75%          Mismatches:      0
Query Match:              99.71%          Indels:          17
DB:                       6               Gaps:            1

US-10-020-095-4 (1-1428) x AX534938 (1-5895)

Qy        1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy       21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy       41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292
```

```
Qy        61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy       101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy       121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy       141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy       161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy       181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy       201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy       221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy       241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy       261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy       281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy       301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy       321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
             |||||||||||||||||||||||||||||||||||||||||||||||||:::||||||||
```

```
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       1218 ---------------------------------------LeuAlaValValGlnPro 1223
                                                     ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4433 ATGGAACTTTGGCTG 4447
```

```
RESULT 5
AX537612
LOCUS       AX537612                4335 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 3 from Patent WO02070738.
ACCESSION   AX537612
VERSION     AX537612.1  GI:25269535
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Schuh,A. and Ouwehand,W.
  TITLE     Diagnosis and treatment of blood disorders
  JOURNAL   Patent: WO 02070738-A 3 12-SEP-2002;
            Schuh, Andre (CA)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..>4335
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57497.1"
                     /db_xref="GI:25269536"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1295 a    838 c    922 g    1280 t
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:          4335
Score:                  7317.50        Matches:         1426
Percent Similarity:     98.75%         Conservative:    1
Best Local Similarity:  98.69%         Mismatches:      1
Query Match:            99.58%         Indels:          17
DB:                     6              Gaps:            1

US-10-020-095-4 (1-1428) x AX537612 (1-4335)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTG 780

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
```

```
Qy        1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy        1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                         ||||||||||||||||||
Db        3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy        1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy        1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy        1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy        1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy        1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy        1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy        1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy        1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy        1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy        1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy        1424 MetGluLeuTrpLeu 1428
               |||||||||||||||
Db        4321 ATGGAACTTTGGCTG 4335
```

```
RESULT 6
AX534936
LOCUS       AX534936                4761 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 3 from Patent WO02070696.
ACCESSION   AX534936
VERSION     AX534936.1  GI:25261478
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 3 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1. .4761
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113. .4450
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57232.1"
                     /db_xref="GI:25261479"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1419 a    913 c    1016 g    1413 t
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:         4761
Score:                  7317.50        Matches:        1426
Percent Similarity:     98.75%         Conservative:   1
Best Local Similarity:  98.69%         Mismatches:     1
Query Match:            99.58%         Indels:         17
DB:                     6              Gaps:           1

US-10-020-095-4 (1-1428) x AX534936 (1-4761)
```

```
Qy           1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy          21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy          41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy          61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy          81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy         101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy         121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532
```

```
Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAAGCCAGTG 892

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              ||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              |||||||||||||||||||||||||||||||||||||||||||:::||||||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4433 ATGGAACTTTGGCTG 4447
```

```
RESULT 7
AX534940
LOCUS       AX534940                5895 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 7 from Patent WO02070696.
ACCESSION   AX534940
VERSION     AX534940.1  GI:25261486
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 7 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1. .5895
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113. .4450
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57234.1"
                     /db_xref="GI:25261487"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
```

```
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1708 a   1174 c   1241 g   1772 t
ORIGIN

Alignment Scores:
Pred. No.:                0               Length:          5895
Score:                    7317.50         Matches:         1426
Percent Similarity:       98.75%          Conservative:    1
Best Local Similarity:    98.69%          Mismatches:      1
Query Match:              99.58%          Indels:          17
DB:                       6               Gaps:            1

US-10-020-095-4 (1-1428) x AX534940 (1-5895)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTAAAAATAAGATAAAG 1792

Qy         561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy         581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy         601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy         621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy         641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy         661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy         681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy         701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
               |||||    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy         721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy         741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy         761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy         781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
               ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy         801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy         821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy         841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy         861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy         881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy         901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy         921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy         941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy         961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy         981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy        1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy        1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy        1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy        1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy        1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy        1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy        1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy        1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy        1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy        1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy        1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy        1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                         ||||||||||||||||||
Db        3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy        1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy        1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952
```

```
Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4433 ATGGAACTTTGGCTG 4447
```

```
RESULT 8
AX534942
LOCUS       AX534942                3535 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 9 from Patent WO02070696.
ACCESSION   AX534942
VERSION     AX534942.1  GI:25261490
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 9 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1. .3535
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113. .3316
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57235.1"
                     /db_xref="GI:25261491"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQVFRI"
BASE COUNT     1108 a    675 c    731 g   1021 t
ORIGIN

Alignment Scores:
Pred. No.:              0            Length:          3535
Score:                  5470.00      Matches:         1062
Percent Similarity:     100.00%      Conservative:    1
Best Local Similarity:  99.91%       Mismatches:      0
Query Match:            74.44%       Indels:          0
DB:                     6            Gaps:            0

US-10-020-095-4 (1-1428) x AX534942 (1-3535)
```

```
Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTG 892
```

```
Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy         1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy         1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy         1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy         1061 LysTyrGln 1063
                |||||||||
Db         3293 AAGTATCAG 3301
```

```
RESULT 9
AX534944
LOCUS       AX534944                3535 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 11 from Patent WO02070696.
ACCESSION   AX534944
VERSION     AX534944.1  GI:25261494
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 11 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1..3535
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113..3316
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57236.1"
                     /db_xref="GI:25261495"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQVFRI"
BASE COUNT     1107 a    676 c    731 g   1021 t
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          3535
Score:                    5461.00        Matches:         1061
Percent Similarity:       99.91%         Conservative:    1
Best Local Similarity:    99.81%         Mismatches:      1
Query Match:              74.32%         Indels:          0
DB:                       6              Gaps:            0

US-10-020-095-4 (1-1428) x AX534944 (1-3535)

Qy            1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy           21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy           41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy           61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy           81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy          101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy          121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy          141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy          161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy          181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy          201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy          221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy          241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy          261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy          281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy          301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy          321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy          341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1192

Qy          361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy          381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy          401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy          421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy          441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy          461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy          481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy          501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy          521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy          541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy          561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy          581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy          601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy          621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy          641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy          661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy          681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy          701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
                ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy          721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy          741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy          761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy          781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
                ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db         2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy          801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy          821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy          841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy          861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy          881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy          901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy          921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy          941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy          961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy          981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy         1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy         1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy         1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy         1061 LysTyrGln 1063
                |||||||||
Db         3293 AAGTATCAG 3301
```

```
RESULT 10
AY083458
LOCUS       AY083458                4735 bp    mRNA    linear   ROD 08-APR-2002
DEFINITION  Mus musculus GPI-anchored alpha-2 macroglobulin-related protein
            mRNA, complete cds.
ACCESSION   AY083458
VERSION     AY083458.1  GI:20070079
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4735)
  AUTHORS   Hashimoto,M., Ichihara,M. and Takahashi,M.
  TITLE     Cloning and characterization of GPI-anchored alpha-2
            macroglobulin-related protein
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4735)
  AUTHORS   Hashimoto,M., Ichihara,M. and Takahashi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2002) Department of Pathology, Nagoya University
            Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya,
            Aichi 466-8550, Japan
FEATURES             Location/Qualifiers
     source          1..4735
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="NIH3T3"
     CDS             343..4671
                     /note="GARP"
                     /codon_start=1
                     /product="GPI-anchored alpha-2 macroglobulin-related
                     protein"
                     /protein_id="AAM00021.1"
                     /db_xref="GI:20070080"
                     /translation="MRSRRLLSAAHLLCLCAVALAAPGSRFLVTAPGIIRPGANVTIG
                     VDLLENSPPQVLVKAQVFKIASNKSRSILEAEGVFHRGHFKTLVLPALPLSSADKIYE
                     LHINGQSENEIVFSNRTRLTFESKSISVLIQTDKAFYKPKQEVKFRVLTLCSDLKPYR
                     TSVDIFIKDPKSNVIQQWFSQKGDLGVVSKTFQLSSNPIFGDWSIQVQVNDQQYYQSF
                     QVLEYVLPKFEVTVQTPLYCSLKSKQLNGSVIAKYTYGKPVKGSLSLTFLPLSFWGKK
                     KNITKSFEINGFANFSFDNYEMKKVMNLKPLTDVSEGSYENVDPSFPGPAEIIATVTE
                     SLTGISRMASTNVFFKQHDYIIEIFDYTTVLKPSLNFTATVKVSRSDGNQLTPEEIEN
                     DLVTVVTQRKNNHPESQRDQEMDYIQTVNYTIPQNGIIKIEFPVMSISGELQLKAYFL
                     DGTSSVTVHSMFTSPSKTYIQLKTRDEYIKVGSPFDLMVSGNRQFKDLSYMVISKGQL
                     VAAGKQSSRTFSLTPEASWAPKACIIAYYIAEDGEIINDILKIPVQLVFENKVKLFWS
                     KPTVKPSDKVSLRISATQSDSLVGIVAVDKSVTLMENSNSITMETMVHELELYNTEYY
                     LGMFMNSFAVFQECGLWVLTDATLIRDSIDEVYDTEEYSERFAEENEANLVDFEDASS
                     VNNVHVRKNFPETWIWLDAYMGSKIYEEFEVTVPDSITSWVASAFVISEDLGFGLTTV
                     PAELQAFQPFFLFLNLPYSVIRGEEFALEVSIVNYLKDTIKVVILIEESDSFDILMTS
                     NDTNGTIYRKTVQVPRDNGVTLVFPIKPTHLGEIPITVTAASPTASDAVTQTIVVKPE
                     GIEKSYSKSVLLDLTDSNVESKQQSMRFSFPPDTVIGSERVQITAIGDILGSSINGLS
                     SLIRMPYGCGEQNMIYFAPNIYILDYLTKQKQLTVNLKEKALSYMRQGYQRELLYQRE
                     DGSFSAFGDIDSSGSTWLSAFVLRCFLEADYYIDIDQDVLHRTYTWLNAHKKFNGEFW
                     EPGRVIHSELQGGTKSPVTLTAYIVTSVLGYKKYQPNIDVQDSIKFLEFEFSRGISDN
                     YTLAIISYALSTVGSPKAEEALNLLMQRSEKEGDTQFWLSSGPALSGSWQPRSVDIEI
                     AAYALLAHTLHHVSEGIPVMRWLIQQRNSLGGFVSTQDTVVALKALSEFSALVHKENT
                     DIQLTVTGPGIPRSIHFRIDSQNLFLLHQEELHALDPITVNVSAHGSGFAICQLNVDY
                     NVKGSGSSKRRRSTENQEVFDLDVIVNNEDDISHLNLNVCTSHLGSERTGMVLMEVNL
                     LSGFSASSDSIPLSETLKKVEYDNGKLNLYLDSVNESQFCVNIPTVRDYKVSNIRDGS
                     VSVMDYYEPRRQAVRSYNTQVKLSSCYLSPDTNCKSHTDGATDSLRRSSSLLVFCSVL
                     LYFVQH"
BASE COUNT     1338 a   1034 c   1087 g   1276 t
ORIGIN

Alignment Scores:
Pred. No.:              0               Length:         4735
Score:                  5360.00         Matches:        1038
Percent Similarity:     83.72%          Conservative:   176
Best Local Similarity:  71.59%          Mismatches:     202
Query Match:            72.95%          Indels:         34
DB:                     10              Gaps:           8
```

```
US-10-020-095-4 (1-1428) x AY083458 (1-4735)

Qy        1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
            |||:::          ||||||:::||||||||||||||||:::|||      ||||||
Db      343 ATGCGCAGTCGGAGGCTCCTGAGCGCGGCCCACCTTCTCTGCTTGTGCGCGGTCGCGCTG 402

Qy       21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
            |||   |||||||||   |||||||||||||||||||||||||||||||||||||
Db      403 GCC---GCCCCAGGGTCCCGCTTTCTGGTGACAGCCCCAGGGATCATCAGACCCGGAGCA 459

Qy       41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
            ||||||||||||||||||:::|||||||||:::   |||   ||||||   ||||||||
Db      460 AATGTGACTATTGGGGTGGATCTCCTGGAAAATAGCCCCCCACAGGTCCTTGTAAAGGCT 519

Qy       61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
            ::::::   |||   |||||||||   :::   |||:::|||||||||||||||||||||
Db      520 CAGGTGTTCAAGATAGCTTCCAACAAATCGAGATCCATCCTGGAAGCAGAAGGAGTCTTT 579

Qy       81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
               :::|||   ||||||||||||   ||||||:::|||||||||:::|||||||||:::
Db      580 CACAGAGGCCATTTCAAGACTCTCGTTCTTCCGGCACTACCTCTGAGCAGTGCAGATAAG 639

Qy      101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
            ||||||||||||      :::      |||:::::::::::||||||:::|||||||||   |||
Db      640 ATTTATGAGCTACACATAAACGGACAATCAGAGAATGAGATCGTATTCTCCAACAGGACA 699

Qy      121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
            ||||||:::||||||:::|||   ||||||||||||   |||||||||||||||   |||
Db      700 CGTTTAACATTTGAGAGCAAGAGTATATCTGTCCTCATTCAGACAGACAAGGCATTCTAC 759

Qy      141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
            |||||||||||||||||||||||||||::::::||||||   ||||||   |||||||||
Db      760 AAGCCGAAGCAGGAAGTGAAGTTTCGGGTGCTCACACTCTGCTCAGATCTCAAGCCTTAT 819

Qy      161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
            :::||||||::::::|||   ||||||||||||||||||||||||:::||||||||||||
Db      820 AGGACATCCGTGGACATTTTCATTAAGGATCCGAAGTCTAATGTGATCCAACAGTGGTTT 879

Qy      181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
            ||||||:::   ||||||||||||:::||||||||||||||||||||||||:::||||||
Db      880 TCTCAGAAAGGTGATCTTGGAGTCGTTTCCAAAACTTTTCAACTATCCTCTAATCCAATA 939

Qy      201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
               |||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db      940 TTTGGTGACTGGTCCATTCAGGTTCAAGTGAATGATCAGCAATATTACCAGTCATTTCAG 999

Qy      221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
            |||   |||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db     1000 GTCCTAGAATATGTGTTACCAAAGTTCGAAGTCACCGTGCAGACCCCGCTGTATTGTTCC 1059

Qy      241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
            :::   ||||||   ||||||||||::::::   ||||||||||||||||||||||||
Db     1060 TTGAAGTCTAAACAGTTAAATGGCTCCGTCATAGCAAAGTATACATACGGGAAGCCAGTG 1119

Qy      261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
            ||||||   ::::::||||||||||||||||||||||||||||||||||||||||||||
Db     1120 AAGGGAAGTCTGTCACTTACATTTTTACCTTTGTCTTTTTGGGGAAAGAAGAAGAATATC 1179

Qy      281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
            ||||||:::|||:::||||||||||   |||||||||||||||::::::   ||||||
Db     1180 ACAAAAAGTTTTGAGATAAATGGATTTGCAAACTTCTCTTTTGATAATTATGAGATGAAA 1239

Qy      301 AsnValMet------------------AspSerSerAsnGlyLeuSerGluTyrLeuAsp 314
               ||||||                  |||   |||   |||      |||   :::|||
Db     1240 AAGGTAATGAATTTGAAACCGCTAACGGATGTCTCGGAAGGGAGTTATGAGAATGTGGAC 1299

Qy      315 LeuSerSerProGlyProValGluIleLeuThrThrValThrGluSerValThrGlyIle 334
               |||   |||||||||   |||||||:::   ||||||||||||||:::|||||||||
Db     1300 CCGTCCTTTCCTGGGCCGGCAGAGATTATAGCTACTGTGACGGAATCACTCACAGGTATC 1359

Qy      335 SerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePhe 354
            ||||||      ||||||||||||||||||||||||||||||||||||||||   |||
```

```
Db   1360 TCAAGAATGGCAAGCACCAATGTGTTTTTCAAGCAACATGATTACATCATTGAGATTTTT 1419

Qy    355 AspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLysValThrArg 374
Db   1420 GATTATACGACTGTCTTGAAACCATCTCTCAACTTCACAGCCACTGTGAAGGTCAGCCGC 1479

Qy    375 AlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIleThrValThr 394
Db   1480 TCTGATGGCAATCAACTAACTCCTGAAGAAATAGAAAATGATTTGGTCACAGTAGTGACC 1539

Qy    395 GlnArg------AsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMet 412
Db   1540 CAGCGGAAAAATAACCACCCTGAG---------------AGTCAGAGAGATCAGGAAATG 1584

Qy    413 GluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGluPhe 432
Db   1585 GATTATATCCAGACCGTAAATTATACTATCCCCAAAATGGAATCATTAAGATTGAATTC 1644

Qy    433 ProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSer 452
Db   1645 CCAGTCATGTCCATTTCCGGCGAGCTACAACTGAAGGCTTATTTTCTTGATGGTACAAGT 1704

Qy    453 SerMetAlaValHisSerLeuPheLysSerProSerLysThrTyrIleGlnLeuLysThr 472
Db   1705 TCTGTGACAGTCCACAGTATGTTCACCTCTCCTAGTAAGACGTACATCCAGCTTAAAACC 1764

Qy    473 ArgAspGluAsnIleLysValGlySerProPheGluLeuValValSerGlyAsnLysArg 492
Db   1765 AGAGATGAATATATAAAGGTGGGGTCACCTTTTGATTTGATGGTTAGTGGCAACAGACAA 1824

Qy    493 LeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGln 512
Db   1825 TTCAAGGACTTAAGCTATATGGTAATATCAAAGGGACAGTTGGTGGCTGCAGGCAAACAA 1884

Qy    513 AsnSerThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIle 532
Db   1885 AGTTCAAGAACTTTCTCTTTAACACCAGAGGCTTCCTGGGCTCCAAAGGCCTGTATAATC 1944

Qy    533 ValTyrTyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGln 552
Db   1945 GCATATTATATTGCAGAGGATGGGGAAATTATAAATGATATTCTAAAAATCCCCGTTCAG 2004

Qy    553 LeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGluProSerGlu 572
Db   2005 CTTGTTTTTGAAAATAAGGTAAAGCTGTTTTGGAGTAAACCTACAGTCAAGCCATCTGAT 2064

Qy    573 LysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaVal 592
Db   2065 AAGGTCTCTCTCAGGATCTCTGCGACACAGTCTGACTCCCTGGTGGGGATTGTAGCTGTT 2124

Qy    593 AspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValValHis 612
Db   2125 GACAAAAGTGTAACATTGATGGAAAACTCCAACAGTATTACAATGGAAACCATGGTCCAT 2184

Qy    613 GluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaVal 632
Db   2185 GAACTGGAACTTTATAACACAGAATATTATCTAGGCATGTTCATGAACTCCTTCGCTGTA 2244

Qy    633 PheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIleAsp 652
Db   2245 TTTCAGGAATGTGGTCTCTGGGTGTTGACAGATGCAACCCTTATAAGAGATAGCATTGAT 2304

Qy    653 GlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluGlyHisIle 672
Db   2305 GAAGTCTACGATACTGAAGAGTATTCTGAACGCTTTGCAGAGGAAAATGAGGCAAACCTG 2364

Qy    673 ValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHisPheProGlu 692
Db   2365 GTTGATTTTGAAGATGCTTCTTCAGTTAACAATGTACATGTCAGGAAGAATTTTCCAGAA 2424

Qy    693 ThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThr 712
Db   2425 ACATGGATTTGGCTAGACGCCTACATGGGATCTAAGATTTATGAAGAATTTGAAGTTACT 2484

Qy    713 ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732
Db   2485 GTACCTGACTCCATCACTTCTTGGGTGGCTTCGGCTTTTGTCATCTCAGAAGACCTGGGT 2544

Qy    733 LeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
Db   2545 TTTGGGTTAACAACCGTTCCAGCAGAGCTGCAAGCCTTCCAACCGTTTTTCCTTTTCCTG 2604

Qy    753 AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772
Db   2605 AATCTTCCGTACTCTGTTATCAGAGGTGAAGAGTTTGCTTTGGAAGTATCCATCGTCAAT 2664

Qy    773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
Db   2665 TATTTGAAAGATACAATTAAGGTTGTGATACTCATTGAGGAAAGTGACAGCTTTGATATT 2724

Qy    793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db   2725 TTAATGACTTCAAATGACACCAATGGCACCATATACCGGAAGACTGTTCAAGTTCCCAGG 2784

Qy    813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db   2785 GATAATGGGGTCACTCTTGTTTTTCCCATCAAGCCCACGCACTTGGGAGAGATTCCCATC 2844

Qy    833 ThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValLys 852
Db   2845 ACAGTGACGGCTGCCTCACCCACTGCCTCTGATGCTGTCACCCAGACAATTGTAGTAAAG 2904

Qy    853 AlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArg 872
Db   2905 CCTGAAGGGATAGAAAAGTCGTATTCGAAATCTGTCTTATTGGATCTGACAGATAGCAAC 2964

Qy    873 LeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySer 892
Db   2965 GTAGAAAGTAAGCAGCAATCTATGAGATTCTCCTTTCCTCCTGATACGGTCATTGGCAGT 3024

Qy    893 GluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAla 912
Db   3025 GAAAGAGTTCAGATCACAGCAATTGGAGATATCCTTGGTTCCTCCATCAATGGCTTATCT 3084

Qy    913 SerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsn 932
Db   3085 TCACTGATCCGGATGCCTTATGGATGTGGTGAACAGAACATGATATATTTGCTCCAAAT 3144

Qy    933 IleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLys 952
Db   3145 ATTTACATTTTGGATTATCTGACTAAACAGAAACAGCTGACAGTTAATTTAAAAGAAAAG 3204

Qy    953 AlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGly 972
Db   3205 GCCCTTTCATATATGAGGCAAGGTTACCAAAGGGAGCTTCTCTATCAGAGGGAAGATGGC 3264

Qy    973 SerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPheVal 992
Db   3265 TCCTTCAGTGCTTTTGGGGACATTGACTCTTCTGGGAGCACTTGGCTGTCAGCATTTGTT 3324

Qy    993 LeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArg 1012
Db   3325 TTAAGATGCTTTCTGGAAGCTGATTACTATATAGATATTGATCAGGATGTGTTACACAGA 3384

Qy   1013 ThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArg 1032
Db   3385 ACATATACTTGGCTTAATGCACATAAGAAATTCAATGGTGAATTTTGGGAGCCAGGAAGA 3444

Qy   1033 ValIleHisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIle 1052
Db   3445 GTGATTCACAGTGAACTTCAAGGTGGCACCAAAAGCCCAGTAACACTTACGGCCTATATT 3504

Qy   1053 ValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIle 1072
Db   3505 GTGACTTCTGTCCTGGGATACAAAAAGTATCAGCCTAATATCGATGTACAAGACTCAATC 3564
```

```
Qy       1073 HisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeuIle 1092
                 |||||||||  ||||||||||||||||||||||||||||||||||||||:::|||
Db       3565 AAGTTTTTGGAATTTGAATTCAGCAGAGGAATTTCGGACAATTATACCTTAGCAATTATA 3624

Qy       1093 ThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeuThr 1112
              :::||||||||||||:::||||||||||||||||||:::|||||||||||||:::|||
Db       3625 TCCTATGCCCTGTCCACAGTCGGGAGCCCTAAAGCAGAGGAGGCTTTGAACTTGCTGATG 3684

Qy       1113 TrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSerLysLeuSer 1132
                 |||:::|||:::||||||       ||||||||:::||||||          ||||||
Db       3685 CAGCGATCAGAAAAGGAAGGAGACACACAGTTCTGGTTGTCATCAGGACCTGCGCTCTCT 3744

Qy       1133 AspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSerHis 1152
                 ||||||||||||||||||:::|||||||||:::|||||||||||||||||:::|||
Db       3745 GGGTCCTGGCAGCCGCGCTCCGTGGACATCGAAATCGCTGCTTATGCACTGCTGGCGCAC 3804

Qy       1153 PheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGlnArg 1172
                 |||         ||||||||||||||||:::|||||||||||    :::||||||
Db       3805 ACTCTGCAC---CATGTGTCTGAGGGAATCCCGGTTATGAGGTGGCTCATCCAGCAAAGA 3861

Qy       1173 AsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSer 1192
              |||||||||||||||||      ||||||||||||||||||||||||||||||||||||
Db       3862 AACAGCCTGGGAGGTTTTGTATCCACTCAGGATACTGTTGTGGCCCTGAAGGCATTATCT 3921

Qy       1193 GluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyPro 1212
              ||||||:::||||||::::::   |||    |||:::||||||:::||||||||||||||
Db       3922 GAATTTTCAGCCCTTGTGCACAAGGAAAATACAGATATCCAACTGACCGTCACAGGGCCC 3981

Qy       1213 SerSerPro--------------------------------------------------- 1215
                    |||
Db       3982 GGCATCCCCAGATCCATACACTTCCGAATTGACTCCCAGAACCTCTTTCTCCTTCACCAG 4041

Qy       1216 SerProLeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPhe 1235
                 |||      :::   |||:::    ||||||:::||||||:::|||    ||||||
Db       4042 GAAGAGCTTCATGCGCTAGATCCCATCACCGTTAATGTTTCTGCGCATGGCTCGGGATTT 4101

Qy       1236 AlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArg 1255
              ||||||||||||||||||||    ||||||||||||    ||||||||||:::||||||
Db       4102 GCTATTTGCCAGCTTAATGTTGACTATAACGTGAAAGGTTCAGGTTCTTCTAAAAGACGG 4161

Qy       1256 ArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLysAsp 1275
              ||||||    :::|||||||||   |||||||||||||||   |||       |||:::|||
Db       4162 AGATCTACCGAAAATCAAGAAGTTTTTGATTTAGACGTCATTGTG---AATAATGAGGAC 4218

Qy       1276 AspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGly 1295
              |||::::::|||::::::||||||||||||||||||       |||      |||:::|||
Db       4219 GACATTAGTCACCTGAATCTGAATGTGTGCACAAGTCACTTGGGTTCAGAGAGGACAGGC 4278

Qy       1296 MetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer 1315
              |||   ||||||||||||||||||||||||||||||          |||::::::|||
Db       4279 ATGGTCCTTATGGAAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATTCAATTCCT 4338

Qy       1316 LeuSerGluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAsp 1335
              ||||||||||||:::||||||||||||||||||:::||||||||||||||||||||||||
Db       4339 CTGAGTGAGACCCTGAAGAAAGTGGAATATGATAACGGGAAACTCAACCTTTATTTAGAT 4398

Qy       1336 SerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSer 1355
              ||||||||||||:::||||||||||||||||||||||    ||||||::::::|||||||
Db       4399 TCTGTAAATGAATCCCAGTTTTGTGTTAATATTCCCACTGTGAGAGACTACAAAGTTTCA 4458

Qy       1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaVal 1375
              |||   :::|||   ||||||||::::::||||||||||||||||||||||||||||
Db       4459 AATATTCGAGATGGTTCAGTATCTGTAATGGACTACTATGAACCAAGGAGACAGGCAGTG 4518

Qy       1376 ArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGly 1395
              ||||||||||||::::::|||||||||||||||||   |||      |||
Db       4519 CGAAGCTACAACACCCAGGTGAAGCTGTCCTCATGTTACCTCAGTCCAGACACC---AAC 4575

Qy       1396 CysArgProCysGluAspGlyAlaSerGlySerHisHisHisSerSerValIlePheIle 1415
              |||:::   |||||||||:::   |||         ||||||    :::    :::
Db       4576 TGCAAGTCACACACGGACGGAGCCACGGACTCCCTTCGACGTTCTTCGAGCCTTCTCGTC 4635

Qy       1416 PheCysPheLysLeuLeuTyrPheMetGlu 1425
```

```
                 ||||||      ||||||||||||::::::
Db       4636 TTTTGCTCCGTCCTTCTGTACTTTGTGCAA 4665

RESULT 11
BC052443
LOCUS       BC052443                5644 bp    mRNA    linear   ROD 19-MAY-2003
DEFINITION  Mus musculus cDNA clone MGC:63463 IMAGE:6406841, complete cds.
ACCESSION   BC052443
VERSION     BC052443.1  GI:30851465
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 5644)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 5644)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
            Thomas L. Casavant.
            Web site: http://genome.uiowa.edu
            Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
            Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
            Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
            Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,
            Casavant,T., Soares,M.B.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series:  Plate:  Row:  Column: 0.
FEATURES             Location/Qualifiers
     source          1..5644
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="MGC:63463 IMAGE:6406841"
                     /tissue_type="Brain, enriched mouse brain 12.5dp"
                     /clone_lib="NIH_BMAP_FO0"
                     /lab_host="DH10B"
```

```
                     /note="Vector: pYX-ASC"
     CDS             105. .4433
                     /codon_start=1
                     /product="Unknown (protein for MGC:63463)"
                     /protein_id="AAH52443.1"
                     /db_xref="GI:30851466"
                     /translation="MRSRRLLSAAHLLCLCAVALAAPGSRFLVTAPGIIRPGANVTIG
                     VDLLENSPPQVLVKAQVFKIASNKSRSILEAEGVFHRGHFKTLVLPALPLSSADKIYE
                     LHINGQSENEIVFSNRTRLTFESKSISVLIQTDKAFYKPKQEVKFRVLTLCSDLKPYR
                     TSVDIFIKDPKSNVIQQWFSQKGDLGVVSKTFQLSSNPIFGDWSIQVQVNDQQYYQSF
                     QVLEYVLPKFEVTVQTPLYCSLKSKQLNGSVIAKYTYGKPVKGSLSLTFLPLSFWGKK
                     KNITKSFEINGFANFSFDNYEMKKVMNLKPLTDVSEGSYENVDPSFPGPAEIIATVTE
                     SLTGISRMASTNVFFKQHDYIIEIFDYTTVLKPSLNFTATVKVSRSDGNQLTPEEIEN
                     DLVTVVTQRKNNHPESQRDQEMDYIQTVNYTIPQNGIIKIEFPVMSISGELQLKAYFL
                     DGTSSVTVHSMFTSPSKTYIQLKTRDEYIKVGSPFDLMVSGNRQFKDLSYMVISKGQL
                     VAAGKQSSRTFSLTPEASWAPKACIIAYYIAEDGEIINDILKIPVQLVFENKVKLFWS
                     KPTVKPSDKVSLRISATQSDSLVGIVAVDKSVTLMENSNSITMETMVHELELYNTEYY
                     LGMFMNSFAVFQECGLWVLTDATLIRDSIDEVYDTEEYSERFAEENEANLVDFEDASS
                     VNNVHVRKNFPETWIWLDAYMGSKIYEEFEVTVPDSITSWVASAFVISEDLGFGLTTV
                     PAELQAFQPFFLFLNLPYSVIRGEEFALEVSIVNYLKDTIKVVILIEESDSFDILMTS
                     NDTNGTIYRKTVQVPRDNGVTLVFPIKPTHLGEIPITVTAASPTASDAVTQTIVVKPE
                     GIEKSYSKSVLLDLTDSNVESKQQSMRFSFPPDTVIGSERVQITAIGDILGSSINGLS
                     SLIRMPYGCGEQNMIYFAPNIYILDYLTKQKQLTVNLKEKALSYMRQGYQRELLYQRE
                     DGSFSAFGDIDSSGSTWLSAFVLRCFLEADYYIDIDQDVLHRTYTWLNAHKKFNGEFW
                     EPGRVIHSELQGGTKSPVTLTAYIVTSVLGYKKYQPNIDVQDSIKFLEFEFSRGISDN
                     YTLAIISYALSTVGSPKAEEALNLLMQRSEKEGDTQFWLSSGPALSGSWQPRSVDIEI
                     AAYALLAHTLHHVSEGIPVMRWLIQQRNSLGGFVSTQDTVVALKALSEFSALVHKENT
                     DIQLTVTGPGIPRSIHFRIDSQNLFLLHQEELHALDPITVNVSAHGSGFAICQLNVDY
                     NVKGSGSSKRRRSTENQEVFDLDVIVNNEDDISHLNLNVCTSHLGSERTGMVLMEVNL
                     LSGFSASSDSIPLSETLKKVEYDNGKLNLYLDSVNESQFCVNIPTVRDYKVSNIRDGS
                     VSVMDYYEPRRQAVRSYNTQVKLSSCYLSPDTNCKSHTDGATDSLRRSSSLLVFCSVL
                     LYFVQH"
BASE COUNT    1556 a   1235 c   1251 g   1602 t
ORIGIN

Alignment Scores:
Pred. No.:              0            Length:          5644
Score:                  5360.00      Matches:         1038
Percent Similarity:     83.72%       Conservative:    176
Best Local Similarity:  71.59%       Mismatches:      202
Query Match:            72.95%       Indels:          34
DB:                     10           Gaps:            8

US-10-020-095-4 (1-1428) x BC052443 (1-5644)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
Db        105 ATGCGCAGTCGGAGGCTCCTGAGCGCGGCCCACCTTCTCTGCTTGTGCGCGGTCGCGCTG 164

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
Db        165 GCC---GCCCCAGGGTCCCGCTTTCTGGTGACAGCCCCAGGGATCATCAGACCCGGAGCA 221

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
Db        222 AATGTGACTATTGGGGTGGATCTCCTGGAAAATAGCCCCCCACAGGTCCTTGTAAAGGCT 281

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
Db        282 CAGGTGTTCAAGATAGCTTCCAACAAATCGAGATCCATCCTGGAAGCAGAAGGAGTCTTT 341

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
Db        342 CACAGAGGCCATTTCAAGACTCTCGTTCTTCCGGCACTACCTCTGAGCAGTGCAGATAAG 401

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
Db        402 ATTTATGAGCTACACATAAACGGACAATCAGAGAATGAGATCGTATTCTCCAACAGGACA 461

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
Db        462 CGTTTAACATTTGAGAGCAAGAGTATATCTGTCCTCATTCAGACAGACAAGGCATTCTAC 521

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
Db        522 AAGCCGAAGCAGGAAGTGAAGTTTCGGGTGCTCACACTCTGCTCAGATCTCAAGCCTTAT 581

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
Db        582 AGGACATCCGTGGACATTTTCATTAAGGATCCGAAGTCTAATGTGATCCAACAGTGGTTT 641

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
Db        642 TCTCAGAAAGGTGATCTTGGAGTCGTTTCCAAAACTTTTCAACTATCCTCTAATCCAATA 701

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
Db        702 TTTGGTGACTGGTCCATTCAGGTTCAAGTGAATGATCAGCAATATTACCAGTCATTTCAG 761

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
Db        762 GTCCTAGAATATGTGTTACCAAAGTTCGAAGTCACCGTGCAGACCCCGCTGTATTGTTCC 821

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
Db        822 TTGAAGTCTAAACAGTTAAATGGCTCCGTCATAGCAAAGTATACATACGGGAAGCCAGTG 881

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
Db        882 AAGGGAAGTCTGTCACTTACATTTTTACCTTTGTCTTTTTGGGGAAAGAAGAAGAATATC 941

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
Db        942 ACAAAAAGTTTTGAGATAAATGGATTTGCAAACTTCTCTTTTGATAATTATGAGATGAAA 1001

Qy        301 AsnValMet------------------AspSerSerAsnGlyLeuSerGluTyrLeuAsp 314
Db       1002 AAGGTAATGAATTTGAAACCGCTAACGGATGTCTCGGAAGGGAGTTATGAGAATGTGGAC 1061

Qy        315 LeuSerSerProGlyProValGluIleLeuThrThrValThrGluSerValThrGlyIle 334
Db       1062 CCGTCCTTTCCTGGGCCGGCAGAGATTATAGCTACTGTGACGGAATCACTCACAGGTATC 1121

Qy        335 SerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePhe 354
Db       1122 TCAAGAATGGCAAGCACCAATGTGTTTTTCAAGCAACATGATTACATCATTGAGATTTTT 1181

Qy        355 AspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLysValThrArg 374
Db       1182 GATTATACGACTGTCTTGAAACCATCTCTCAACTTCACAGCCACTGTGAAGGTCAGCCGC 1241

Qy        375 AlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIleThrValThr 394
Db       1242 TCTGATGGCAATCAACTAACTCCTGAAGAAATAGAAAATGATTTGGTCACAGTAGTGACC 1301

Qy        395 GlnArg------AsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMet 412
Db       1302 CAGCGGAAAAATAACCACCCTGAG---------------AGTCAGAGAGATCAGGAAATG 1346

Qy        413 GluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGluPhe 432
Db       1347 GATTATATCCAGACCGTAAATTATACTATCCCCCAAAATGGAATCATTAAGATTGAATTC 1406

Qy        433 ProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSer 452
Db       1407 CCAGTCATGTCCATTTCCGGCGAGCTACAACTGAAGGCTTATTTTCTTGATGGTACAAGT 1466

Qy        453 SerMetAlaValHisSerLeuPheLysSerProSerLysThrTyrIleGlnLeuLysThr 472
Db       1467 TCTGTGACAGTCCACAGTATGTTCACCTCTCCTAGTAAGACGTACATCCAGCTTAAAACC 1526

Qy        473 ArgAspGluAsnIleLysValGlySerProPheGluLeuValValSerGlyAsnLysArg 492
Db       1527 AGAGATGAATATATAAAGGTGGGGTCACCTTTTGATTTGATGGTTAGTGGCAACAGACAA 1586

Qy        493 LeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGln 512
```

```
                 |||:::|||||||||||||||:::|||:::|||||||||||||||      |||||||||
Db       1587 TTCAAGGACTTAAGCTATATGGTAATATCAAAGGGACAGTTGGTGGCTGCAGGCAAACAA 1646

Qy        513 AsnSerThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIle 532
              :::|||     ||||||||||||||||||  ||||||    |||||||||||:::|||
Db       1647 AGTTCAAGAACTTTCTCTTTAACACCAGAGGCTTCCTGGGCTCCAAAGGCCTGTATAATC 1706

Qy        533 ValTyrTyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGln 552
              |||||||||    :::|||||||||||||||:::|||:::||||||||||||||||||
Db       1707 GCATATTATATTGCAGAGGATGGGGAAATTATAAATGATATTCTAAAAATCCCCGTTCAG 1766

Qy        553 LeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGluProSerGlu 572
              |||||||||:::||||||:::|||||||:::|||||||||          :::||||||:::
Db       1767 CTTGTTTTTGAAAATAAGGTAAAGCTGTTTTGGAGTAAACCTACAGTCAAGCCATCTGAT 1826

Qy        573 LysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaVal 592
              ||||||||||||||||||||||    ||||||   ||||||:::||||||||||||||||||
Db       1827 AAGGTCTCTCTCAGGATCTCTGCGACACAGTCTGACTCCCTGGTGGGGATTGTAGCTGTT 1886

Qy        593 AspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValValHis 612
              ||||||||||||   ||||||      ||||||   ||||||||||||     :::||||||
Db       1887 GACAAAAGTGTAACATTGATGGAAAACTCCAACAGTATTACAATGGAAACCATGGTCCAT 1946

Qy        613 GluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaVal 632
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db       1947 GAACTGGAACTTTATAACACAGAATATTATCTAGGCATGTTCATGAACTCCTTCGCTGTA 2006

Qy        633 PheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIleAsp 652
              ||||||||||||||||||||||||||||||||||||    |||    :::|||    ||||||
Db       2007 TTTCAGGAATGTGGTCTCTGGGTGTTGACAGATGCAACCCTTATAAGAGATAGCATTGAT 2066

Qy        653 GlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluGlyHisIle 672
                 |||||||||      ||||||:::|||||||||    ||||||||||||      ::::::
Db       2067 GAAGTCTACGATACTGAAGAGTATTCTGAACGCTTTGCAGAGGAAAATGAGGCAAACCTG 2126

Qy        673 ValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHisPheProGlu 692
              ||||||      |||    |||       ::::::    ||||||||||||:::|||||||||
Db       2127 GTTGATTTTGAAGATGCTTCTTCAGTTAACAATGTACATGTCAGGAAGAATTTCCAGAA 2186

Qy        693 ThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThr 712
              ||||||||||||||||||       ||||||    :::||||||:::|||||||||||||||
Db       2187 ACATGGATTTGGCTAGACGCCTACATGGGATCTAAGATTTATGAAGAATTTGAAGTTACT 2246

Qy        713 ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732
              ||||||||||||||||||||||||:::   ||||||||||||||||||||||||||||||
Db       2247 GTACCTGACTCCATCACTTCTTGGGTGGCTTCGGCTTTTGTCATCTCAGAAGACCTGGGT 2306

Qy        733 LeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
                 ||||||||||||    |||   ||||||||||||||||||||||||||:::||||||
Db       2307 TTTGGGTTAACAACCGTTCCAGCAGAGCTGCAAGCCTTCCAACCGTTTTTCCTTTTCCTG 2366

Qy        753 AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772
              |||||||||||||||||||||||||||||||||||||||||||||||:::::::|||   |||
Db       2367 AATCTTCCGTACTCTGTTATCAGAGGTGAAGAGTTTGCTTTGGAAGTATCCATCGTCAAT 2426

Qy        773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
              ||||||||||||       :::|||    ::::::||||||:::||||||    |||||||||
Db       2427 TATTTGAAAGATACAATTAAGGTTGTGATACTCATTGAGGAAAGTGACAGCTTTGATATT 2486

Qy        793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
              ||||||||||||::::::    |||    |||   :::::::::|||:::       ||||||
Db       2487 TTAATGACTTCAAATGACACCAATGGCACCATATACCGGAAGACTGTTCAAGTTCCCAGG 2546

Qy        813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
              ::::::|||   |||::::::|||||||||:::|||||||||||||||||||||||||||
Db       2547 GATAATGGGGTCACTCTTGTTTTTCCCATCAAGCCCACGCACTTGGGAGAGATTCCCATC 2606

Qy        833 ThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValLys 852
              ||||||||||||   ||||||||||||||||||||||||||||||||     |||:::||||||
Db       2607 ACAGTGACGGCTGCCTCACCCACTGCCTCTGATGCTGTCACCCAGACAATTGTAGTAAAG 2666

Qy        853 AlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArg 872
              ||||||||||||||||||||||||:::|||:::|||||||||||||||||||:::

Db       2667 CCTGAAGGGATAGAAAAGTCGTATTCGAAATCTGTCTTATTGGATCTGACAGATAGCAAC 2726

Qy        873 LeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySer 892
              ::::::|||      :::::::::   ||||||||||||||:::||||||   ||||||
Db       2727 GTAGAAAGTAAGCAGCAATCTATGAGATTCTCCTTTCCTCCTGATACGGTCATTGGCAGT 2786

Qy        893 GluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAla 912
              |||||||||||||||||||||||||||:::||||||     ||||||||||||||||:::
Db       2787 GAAAGAGTTCAGATCACAGCAATTGGAGATATCCTTGGTTCCTCCATCAATGGCTTATCT 2846

Qy        913 SerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsn 932
              ||||||||||||||||||||||||||||||||||||||||||||    ||||||||||||
Db       2847 TCACTGATCCGGATGCCTTATGGATGTGGTGAACAGAACATGATATATTTTGCTCCAAAT 2906

Qy        933 IleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLys 952
              ||||||||||||||||||||||||||||:::||||||||||||   |||||||||||||||
Db       2907 ATTTACATTTTGGATTATCTGACTAAACAGAAACAGCTGACAGTTAATTTAAAAGAAAAG 2966

Qy        953 AlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGly 972
              |||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2967 GCCCTTTCATATATGAGGCAAGGTTACCAAAGGGAGCTTCTCTATCAGAGGGAAGATGGC 3026

Qy        973 SerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPheVal 992
              ||||||||||||||||||:::   |||   |||||||||||||||||||||||||||||
Db       3027 TCCTTCAGTGCTTTTGGGGACATTGACTCTTCTGGGAGCACTTGGCTGTCAGCATTTGTT 3086

Qy        993 LeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArg 1012
              ||||||||||||||||||||||||    ||||||||||||||||:::||||||||||||
Db       3087 TTAAGATGCTTTCTGGAAGCTGATTACTATATAGATATTGATCAGGATGTGTTACACAGA 3146

Qy       1013 ThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArg 1032
              ||||||||||||||||||     |||:::|||   ||||||||||||||:::||||||||||
Db       3147 ACATATACTTGGCTTAATGCACATAAGAAATTCAATGGTGAATTTTGGAGCCAGGAAGA 3206

Qy       1033 ValIleHisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIle 1052
              ||||||||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db       3207 GTGATTCACAGTGAACTTCAAGGTGGCACCAAAAGCCCAGTAACACTTACGGCCTATATT 3266

Qy       1053 ValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIle 1072
              ||||||||:::|||||||||:::|||||||||||||||||||||||||||:::||||||
Db       3267 GTGACTTCTGTCCTGGGATACAAAAAGTATCAGCCTAATATCGATGTACAAGACTCAATC 3326

Qy       1073 HisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeuIle 1092
                 |||||||||   ||||||||||||||||||||||||||||||||||||:::|||
Db       3327 AAGTTTTTGGAATTTGAATTCAGCAGAGGAATTTCGGACAATTATACCTTAGCAATTATA 3386

Qy       1093 ThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeuThr 1112
              :::|||||||||||||||||||||||||:::||||||||||||||||||:::|||
Db       3387 TCCTATGCCCTGTCCACAGTCGGGAGCCCTAAAGCAGAGGAGGCTTTGAACTTGCTGATG 3446

Qy       1113 TrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSerLysLeuSer 1132
                 |||:::|||:::||||||   ||||||||||||:::|||||||      ||||||
Db       3447 CAGCGATCAGAAAAGGAAGGAGACACACAGTTCTGGTTGTCATCAGGACCTGCGCTCTCT 3506

Qy       1133 AspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSerHis 1152
                 ||||||||||||||||||:::||||||||:::||||||||||||||||:::|||
Db       3507 GGGTCCTGGCAGCCGCGCTCCGTGGACATCGAAATCGCTGCTTATGCACTGCTGGCGCAC 3566

Qy       1153 PheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGlnArg 1172
                 |||         ||||||||||||:::|||||||||||||     :::||||||
Db       3567 ACTCTGCAC---CATGTGTCTGAGGGAATCCCGGTTATGAGGTGGCTCATCCAGCAAAGA 3623

Qy       1173 AsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSer 1192
              ||||||||||||||||   ||||||||||||||     |||||||||||||||||||
Db       3624 AACAGCCTGGGAGGTTTTGTATCCACTCAGGATACTGTTGTGGCCCTGAAGGCATTATCT 3683

Qy       1193 GluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyPro 1212
              ||||||:::||||||::::::    |||    |||:::||||||:::||||||||||||
Db       3684 GAATTTTCAGCCCTTGTGCACAAGGAAAATACAGATATCCAACTGACCGTCACAGGGCCC 3743

Qy       1213 SerSerPro--------------------------------------------------- 1215
                    |||
Db       3744 GGCATCCCCAGATCCATACACTTCCGAATTGACTCCCAGAACCTCTTTCTCCTTCACCAG 3803
```

```
Qy       1216 SerProLeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPhe 1235
                |||      :::   |||:::   ||||||:::||||||:::|||   ||||||
Db       3804 GAAGAGCTTCATGCGCTAGATCCCATCACCGTTAATGTTTCTGCGCATGGCTCGGGATTT 3863

Qy       1236 AlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArg 1255
              ||||||||||||||||||||   ||||||||||||   ||||||||||||:::||||||
Db       3864 GCTATTTGCCAGCTTAATGTTGACTATAACGTGAAAGGTTCAGGTTCTTCTAAAAGACGG 3923

Qy       1256 ArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLysAsp 1275
              ||||||   :::|||||||||   |||||||||||||||   |||      |||:::|||
Db       3924 AGATCTACCGAAAATCAAGAAGTTTTTGATTTAGACGTCATTGTG---AATAATGAGGAC 3980

Qy       1276 AspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGly 1295
              |||::::::|||:::||||||||||||||||||   |||      |||:::|||
Db       3981 GACATTAGTCACCTGAATCTGAATGTGTGCACAAGTCACTTGGGTCAGAGAGGACAGGC 4040

Qy       1296 MetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer 1315
              |||   ||||||||||||||||||||||||||||||         |||::::::|||
Db       4041 ATGGTCCTTATGGAAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATTCAATTCCT 4100

Qy       1316 LeuSerGluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAsp 1335
              ||||||||||||:::||||||||||||||||||:::||||||||||||||||||||||||
Db       4101 CTGAGTGAGACCCTGAAGAAAGTGGAATATGATAACGGGAAACTCAACCTTTATTTAGAT 4160

Qy       1336 SerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSer 1355
              ||||||||||||:::||||||||||||||||||||||   ||||||::::::|||||||||
Db       4161 TCTGTAAATGAATCCCAGTTTTGTGTTAATATTCCCACTGTGAGAGACTACAAAGTTTCA 4220

Qy       1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaVal 1375
              |||   :::|||   ||||||||::::::||||||||||||||||||||||||||||||
Db       4221 AATATTCGAGATGGTTCAGTATCTGTAATGGACTACTATGAACCAAGGAGACAGGCAGTG 4280

Qy       1376 ArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGly 1395
              ||||||||||||::::::||||||||||||||||||   |||      |||
Db       4281 CGAAGCTACAACACCCAGGTGAAGCTGTCCTCATGTTACCTCAGTCCAGACACC---AAC 4337

Qy       1396 CysArgProCysGluAspGlyAlaSerGlySerHisHisHisSerSerValIlePheIle 1415
              |||:::      |||||||||:::   |||         ||||||   :::   :::
Db       4338 TGCAAGTCACACACGGACGGAGCCACGGACTCCCTTCGACGTTCTTCGAGCCTTCTCGTC 4397

Qy       1416 PheCysPheLysLeuLeuTyrPheMetGlu 1425
              ||||||      ||||||||||||::::::
Db       4398 TTTGCTCCGTCCTTCTGTACTTTGTGCAA 4427

RESULT 12
HSM805581
LOCUS       HSM805581               2938 bp    mRNA    linear   PRI 12-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp762L1111 (from clone DKFZp762L1111).
ACCESSION   AL834478
VERSION     AL834478.1  GI:21740252
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2938)
  AUTHORS   Blum,H., Bauersachs,S., Mewes,H.W., Weil,B. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by LMU (Ludwig Maximilians University,
            Munich/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp762L1111) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1..2938
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="15"
                     /clone="DKFZp762L1111"
                     /tissue_type="melanoma (MeWo cell line)"
                     /clone_lib="762 (synonym: hmel2). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     gene            1..2938
                     /gene="DKFZp762L1111"
     CDS             <1..2565
                     /gene="DKFZp762L1111"
                     /note="similarity to alpha-2-macroglobulin (Limulus sp.)"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAD39137.1"
                     /db_xref="GI:21740253"
                     /translation="VDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQE
                     CGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGSSPHVRKHFPET
                     WIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIF
                     LNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLL
                     VPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAITQMILVKAEGIEKSYSQSILLD
                     LTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQN
                     MINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPS
                     GSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG
                     NKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSV
                     GSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQ
                     TSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSP
                     SPVKFLIDTHNRLLLQTAELAVVQPTAVNISANGFGFAICQLNVVYNVKASGSSRRRR
                     SIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAI
                     SLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRR
                     QAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYFMELWL"
     polyA_signal    2843..2848
                     /gene="DKFZp762L1111"
     polyA_site      2859
                     /gene="DKFZp762L1111"
BASE COUNT      887 a    529 c    620 g    902 t
ORIGIN

Alignment Scores:
Pred. No.:              5.54e-274      Length:        2938
Score:                  4305.50        Matches:       834
Percent Similarity:     97.78%         Conservative:  1
Best Local Similarity:  97.66%         Mismatches:    2
Query Match:            58.59%         Indels:        17
DB:                     9              Gaps:          1

US-10-020-095-4 (1-1428) x HSM805581 (1-2938)

Qy        592 ValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValVal 611
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTGACAAAAGTGTGAATCTGATGAATGCCTCTAATGATATTACAATGGAAAATGTGGTC 60

Qy        612 HisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAla 631
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGAGTTGGAACTTTATAACACAGGATATTATTTAGGCATGTTCATGAATTCTTTTGCA 120

Qy        632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCTTTCAGGAATGTGGACTCTGGGTATTGACAGATGCAAACCTCACGAAGGATTATATT 180

Qy        652 AspGlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluGlyHis 671
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGGTGTTTATGACAATGCAGAATATGCTGAGAGGTTTATGGAGGAAAATGAAGGACAT 240

Qy        672 IleValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHisPhePro 691
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTGTAGATATTCATGACTTTTCTTTGGGTAGCAGTCCACATGTCCGAAAGCATTTTCCA 300

Qy        692 GluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluVal 711
              ||||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db        301 GAGACTTGGATTTGGCTAGACACCAACATGGGTTCCAGGATTTACCAAGAATTTGAAGTA 360
```

```
Qy    712 ThrValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeu 731
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ACTGTACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTTGTGATCTCTGAGGACCTG 420

Qy    732 GlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePhe 751
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTCCAACCATTTTTCATTTTT 480

Qy    752 LeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePhe 771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGAATTTGCTTTGGAAATAACTATATTC 540

Qy    772 AsnTyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAsp 791
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AATTATTTGAAAGATGCCACTGAGGTTAAGGTAATCATTGAGAAAAGTGACAAATTTGAT 600

Qy    792 IleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValPro 811
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ATTCTAATGACTTCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCC 660

Qy    812 SerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIlePro 831
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 AGTGAGGATGGGGCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCT 720

Qy    832 IleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuVal 851
          |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db    721 ATCACAGTCACAGCTCTTTCACCCACTGCTTCTGATGCTATCACCCAGATGATTTTAGTA 780

Qy    852 LysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsn 871
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 AAGGCTGAAGGAATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAAT 840

Qy    872 ArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGly 891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AGGCTACAGAGTACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGC 900

Qy    892 SerGluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeu 911
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 AGTGAAAGAGTTCAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTA 960

Qy    912 AlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaPro 931
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCCTCATTGATTCGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCA 1020

Qy    932 AsnIleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGlu 951
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 AATATTTACATTTTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAA 1080

Qy    952 LysAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAsp 971
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAAGCTCTTTCATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGAT 1140

Qy    972 GlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPhe 991
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GGCTCTTTCAGTGCTTTTGGGAATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTT 1200

Qy    992 ValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHis 1011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 GTTTTAAGATGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACAC 1260

Qy   1012 ArgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGly 1031
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGAACATACACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGA 1320

Qy   1032 ArgValIleHisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyr 1051
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 AGAGTGATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTAT 1380

Qy   1052 IleValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSer 1071
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ATTGTAACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCT 1440

Qy   1072 IleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeu 1091
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ATCCATTTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCTAGCCCTT 1500

Qy   1092 IleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeu 1111
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 ATAACTTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAATATGCTG 1560

Qy   1112 ThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSerLysLeu 1131
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 ACTTGGAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTT 1620

Qy   1132 SerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSer 1151
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCA 1680

Qy   1152 HisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGln 1171
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 CACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAA 1740

Qy   1172 ArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeu 1191
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 AGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAAGGCTCTG 1800

Qy   1192 SerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGly 1211
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTGACGGGG 1860

Qy   1212 ProSerSerProSerPro------------------------------------------ 1217
          ||||||||||||||||||
Db   1861 CCTAGCTCACCAAGTCCTGTAAAGTTTCTGATTGACACACACAACCGCTTACTCCTTCAG 1920

Qy   1218 ---------LeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGly 1234
                   ||||||||||||||||||    ||||||||||||||||||||||||||||||
Db   1921 ACAGCAGAGCTTGCTGTGGTACAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTGGA 1980

Qy   1235 PheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArg 1254
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 TTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGA 2040

Qy   1255 ArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLys 1274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 CGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAA 2100

Qy   1275 AspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSer 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGT 2160

Qy   1295 GlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIle 1314
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATT 2220

Qy   1315 SerLeuSerGluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeu 1334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 TCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTA 2280

Qy   1335 AspSerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysVal 1354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTT 2340

Qy   1355 SerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAla 1374
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 TCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCG 2400

Qy   1375 ValArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGln 1394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 GTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAG 2460

Qy   1395 GlyCysArgProCysGluAspGlyAlaSerGlySerHisHisHisSerSerValIlePhe 1414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTT 2520

Qy   1415 IlePheCysPheLysLeuLeuTyrPheMetGluLeuTrpLeu 1428
          ||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2521 ATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTG 2562

RESULT 13
AK095888
LOCUS       AK095888                2273 bp    mRNA    linear   PRI 15-JUL-2002
DEFINITION  Homo sapiens cDNA FLJ38569 fis, clone HCHON2006459.
ACCESSION   AK095888
VERSION     AK095888.1  GI:21755236
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
            Arita,M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B.,
            Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and
            Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2273)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1. .2273
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HCHON2006459"
                     /cell_type="chondrocytes (HC)"
                     /clone_lib="HCHON2"
                     /note="cloning vector: pME18SFL3~primary culture,
                     chondrocytes"
     CDS             66. .2063
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC04642.1"
                     /db_xref="GI:21755237"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNLFGTQEAL"
BASE COUNT      712 a     431 c     462 g     668 t
ORIGIN

Alignment Scores:
Pred. No.:              6.9e-212        Length:         2273
Score:                  3360.00         Matches:        657
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
```

```
Query Match:            45.73%          Indels:         0
DB:                     9               Gaps:           0

US-10-020-095-4 (1-1428) x AK095888 (1-2273)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 125

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 185

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 245

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 305

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 365

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 425

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 485

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 545

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 605

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 665

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAG 725

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 785

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 845

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        846 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 905

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 965

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        966 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1025

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1085
```

```
Qy      341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1086 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1145

Qy      361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1146 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1205

Qy      381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1206 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1265

Qy      401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1266 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1325

Qy      421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1326 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1385

Qy      441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1386 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1445

Qy      461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1446 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1505

Qy      481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1506 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1565

Qy      501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1566 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1625

Qy      521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1626 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1685

Qy      541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1686 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1745

Qy      561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1746 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1805

Qy      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1806 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1865

Qy      601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1866 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1925

Qy      621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1926 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1985

Qy      641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsn 657
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1986 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAAT 2036
```

RESULT 14
AX054953

```
LOCUS       AX054953                4527 bp    DNA     linear   PAT 13-JAN-2001
DEFINITION  Sequence 68 from Patent WO0073328.
ACCESSION   AX054953
VERSION     AX054953.1  GI:12228337
KEYWORDS    .
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1
  AUTHORS   van Criekinge,W., Roelens,I., Bogaert,T. and Verwaerde,P.
  TITLE     Unc-5 constructs and screening methods
  JOURNAL   Patent: WO 0073328-A 68 07-DEC-2000;
            Devgen NV (BE)
FEATURES             Location/Qualifiers
     source          1..4527
                     /organism="Caenorhabditis elegans"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:6239"
BASE COUNT     1313 a    948 c   1077 g   1189 t
ORIGIN

Alignment Scores:
Pred. No.:              2.27e-112     Length:        4527
Score:                  1854.50       Matches:       476
Percent Similarity:     51.23%        Conservative:  292
Best Local Similarity:  31.75%        Mismatches:    562
Query Match:            25.24%        Indels:        169
DB:                     6             Gaps:          44

US-10-020-095-4 (1-1428) x AX054953 (1-4527)

Qy       15 ValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThrAlaProGly 34
            |||   |||   |||         |||   |||         :::::::::   ||||||
Db       82 GTGTCAACAACCGCGGCGCCAGTTAAACCAGCC---ACTTATATGCTTGTCGCCCCAGCA 138

Qy       35 IleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHisCysProSer 54
            ::::::||||||            ::::::   :::   ||||||:::
Db      139 GTAGTGCGACCCGATCAACCGTTCTCCGTTTGCATGAACCTTCTGAAACAAGCTACGGAT 198

Qy       55 GlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeu 74
            :::      :::            ::::::|||         |||   |||::::::
Db      199 GAAGACATGATTGTTCGGATTGAAGTGAGAACTGAACGGAACGAGACGATTGCTGCTCGA 258

Qy       75 GluAlaGluGlyValPheGluLysGlySerPheLysThrLeuThrLeuProSerLeuPro 94
                        :::   :::   |||         :::|||::::::|||         :::|||
Db      259 GTTATAAGCAATTTG---AAGCCGGGAATCGCACAGACTGTCAGCTTGAGTGAGATGCCT 315

Qy       95 LeuAsnSerAlaAsp------GluIleTyrGluLeuArgValThrGlyArgThrGlnAsp 112
                  |||               :::   |||:::|||   :::   |||   |||   :::
Db      316 GCTCAATCACTAACTCCTAGGCAAAGCTATAAGCTCTACATCCGCGGGGAGACTCTTAAC 375

Qy      113 ---GluIleLeuPheSerAsnSerThrArgLeuSerPheGluThrLysArgIleSerVal 131
               |||::::::|||   |||               |||   ::::::   |||   :::||||||
Db      376 GCAGAACTCATATTTGAAAACGAAAATGAGTTAAAGTACGATCAAAAAGCTTTATCAGTT 435

Qy      132 PheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluValLysPheArgIleVal 151
            |||||||||||||||:::|||:::|||:::|||         |||::::::|||   :::
Db      436 TTCATCCAAACAGATCGAGCAATCTATCGGCCTGCTTCTCTTGTTCGATACCGGGCAATT 495

Qy      152 ThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIleLeuIleLysAspPro 171
               :::   ||||||   |||||||||      :::      |||   |||   ||||||
Db      496 GTAGTGAAATCGGACTTGAAACCTTACGTCGGAAACGCTACTATTAAAATTTTCGATCCA 555

Qy      172 LysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAspLeuGlyValIleSerLys 191
                  |||||||||   |||   :::               |||   ||||||   |||
Db      556 AGCAGAAATCTGATTTCACAGACGATTGGTGTAACTTTGGATAGAGGAGTATACAGCGGC 615

Qy      192 ThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSerIleGlnValGln------ 209
                  ||||||:::         :::||||||||||||   |||:::|||:::
Db      616 GAGCTCCAACTTGCCGAAGAGACTCTCCTTGGAGATTGGTTTATCGAGGTGGAAACCTCG 675

Qy      210 ------ValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrValLeuPro 227
                  |||   ||||||               ||||||   |||      ||||||||||||
Db      676 AATGGAGTCCAGGACAAGTCG---------TCATTTACGGTAGATACATACGTGCTTCCC 726

Qy      228 LysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSerLysHisLeuAsn 247
            ||||||||||||   ::::::|||      :::   ::::::|||         |||:::
Db      727 AAATTTGAAGTGAATATCAAGACGTCTTCGTTTATAACTATTAACGAT---GATTTGTCA 783
```

```
Qy    248 GlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAspValThrLeuThr 267
Db    784 GTGTTCGTCGATGCAAAGTACACATATGGAAAAGGCGTTGCTGGAAAGGCCAAAGTATCT 843
Qy    268 Phe---LeuProLeuSerPheTrp--------------------------------Gly 275
Db    844 CTAGAGCTTCCATGGCACAGATGGCATGCGATGGTTCCCACTATCATTGACGAAAATGGA 903
Qy    276 LysLysLysAsn-----------IleThrLysThrPheLysIleAsn------GlySer 289
Db    904 GTCAAGAAGGAAGAGGAGCTGATGGTCGAGCGTACTGTAAAGCTGAACAGACAAGGAGAA 963
Qy    290 AlaAsnPheSerPheAsnAspGluGluMetLysAsnValMetAspSerSerAsnGlyLeu 309
Db    964 GCGGCAGTTGTTTTCTCGAATGATGAGCTCAAGAGG----------------------- 999
Qy    310 SerGluTyrLeuAspLeuSerSerProGlyProValGluIleLeuThrThrValThrGlu 329
Db   1000 CACAAGTTGCTACATGAATGGGGAGGAGGTTCAATTCGAATTGTTGCCTCAGTCACTGAA 1059
Qy    330 SerValThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyr 349
Db   1060 GACATCACTGAAATCGAAAGAAACGCTACTCATCAAATTTCAACGTTCCGTGAGGAGGTC 1119
Qy    350 IleIleGluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThr 369
Db   1120 AAACTTGATGTTGAAAAGCAGGGAGACACCTTCAAGCCAGGACTCACTTATAATGTTGTA 1179
Qy    370 ValLysValThrArgAlaAspGlyAsn------GlnLeuThrLeuGluGluArgArgAsn 387
Db   1180 GTTGCACTGAAACAAATGGATGACACTCCCGTGAAGGCAACTTTGCCAAAGAGA------ 1233
Qy    388 AsnValValIleThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSer 407
Db   1234 ------GTTCAGGTATCAACATTCTACAACTATCCA------TACAATCATGACACTTCA 1281
Qy    408 GlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThr 427
Db   1282 TCTCTGCAAGAGGAAAAAGAAACAAAAATTGTGGAAGTTGACGCTCACGGCACCTCGGTT 1341
Qy    428 PheLysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPhe 447
Db   1342 CTCACCCTCCAACCACCAATCAACTGCACTTCTGCTCGCATTGAGGCTCAT---TATGAT 1398
Qy    448 LeuGlySerLysSerSerMetAlaValHisSerLeuPheLys------------------ 461
Db   1399 ATTGGTGGCAAGGACAACTTCACAGCCACCCCAATATATTCCTCGCTATATGTCGAGGCA 1458
Qy    462 ------SerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsn---IleLys 478
Db   1459 GCCGTGAGCCCAACGAAGAGTTTCCTGCAGCTTTTGGCTGATAATGAGGGAGCTGTTGAT 1518
Qy    479 ValGlySerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyr 498
Db   1519 GTAGGAAAGAGTCTTAGCTTCTCGTTGAAGGCTACCCAGCCGCTGTCGACAATTACCTAT 1578
Qy    499 MetValValSerArgGlyGlnLeuVal---------------AlaValGlyLysGlnAsn 513
Db   1579 CAGGTCATGTCTCGTTCGAACATTGTGGTGTCTCAACAAATGACTGTTAATTCCGAACAT 1638
Qy    514 SerThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleVal 533
Db   1639 GCGACCATT---AGCTTCCCAGCAACGGCTAACATGGCCCCAAAGTCACGGTTGATCGTC 1695
Qy    534 Tyr---TyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGln 552
Db   1696 TATGCCATTATTGAGAGTAGTCAAGAAGTTCTTGTTGATGCACTTGACTTTAAAGTAGAA 1755
Qy    553 LeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGluProSerGlu 572
Db   1756 GGAATATTCCAAAACCAAGTGGCACTATCAATTGACAAACAAGCAGTCGAGCCAGGACAG 1815
Qy    573 LysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaVal 592
Db   1816 AATGTGAAGTTTAAGGTCACA---TCTGACAAGAACTCTTTTGTTGGGCTTCTTGTGGTA 1872
Qy    593 AspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValValHis 612
Db   1873 GATCAGAGTGTTCTCTTGTTAAAAACTGGCAATGATATCACTAGAGAAAAAGTTGAGCAA 1932
Qy    613 GluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPhe----------------- 626
Db   1933 GACTTGGAGAATTATGACTCTAATAATGTTGGAGGAGGATTCGGGGGTCCACGGCCATGG 1992
Qy    627 ---------------------------------------------------------Met 627
Db   1993 GAGGCAATTGATAGAAAGAAGAGGTCCATTTGGAGACCTTGGTGGGGAATTGGAGGGAGT 2052
Qy    628 AsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThr 647
Db   2053 GATGCTCAATCAATCTTTTCTAATGCTGGGCTTGTAGTTTGACCGACGCCCTCTTGTAC 2112
Qy    648 LysAsp---------------------------TyrIleAspGlyValTyrAspAsnAla 658
Db   2113 CGCGAGCCTCAACGTGAATTTATGTCAGTAATGATGATGGATGGTGCTCCAGGAATGGCT 2172
Qy    659 GluTyrAlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPhe 678
Db   2173 GAAGCCGCA------TTCGCCGCTCCGCCAATGGGC----------------------- 2202
Qy    679 SerLeuGlySerSer---------ProHisValArgLysHisPheProGluThrTrpIle 695
Db   2203 ------GGGTCGTCGCCTCCACCGCCGACAGTTCGAAAGTTCTTCCCACACACCTGGATT 2256
Qy    696 TrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAsp 715
Db   2257 TGGTCGGATCTCAAC---TCTACTAGTGGAGAAGTTGAGATGGAAATTGAAGCTCCGGAC 2313
Qy    716 SerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeu 735
Db   2314 ACCATCACCTCGTGGGTCGCCTCCACTTTTGCAATCAATGAGGAAAATGGTCTCGGTGTG 2373
Qy    736 ThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuPro 755
Db   2374 GCTCCAACGACTTCAAAACTTCGCGTGTTTCGTCCATTTTTTATTCAACTCAATCTTCCA 2433
Qy    756 TyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLys 775
Db   2434 TATGCTGTCCGTCGTGGAGAAAAGTTTGCACTTTTGGTGCTTGTTTTCAATTATATGGAA 2493
Qy    776 AspAlaThrGluValLysValIleIle------GluLysSerAspLysPheAspIleLeu 793
Db   2494 AAGGAGCAAGATGTAACAGTGACTCTGAAGTATGATAAAGACTCTGGTTACGATCTCCTG 2553
Qy    794 MetThrSerSerGluIle---------AsnAlaThrGlyHisGlnGlnThrLeuLeuVal 810
Db   2554 AAGAAAGATGGAACTGTAGTAAGGCGTGACGAGGTTGGCCAACAAAATGTTCGTATTGTG 2613
Qy    811 ProSerGluAspGlyAlaThr---------ValLeuPheProIleArgProThrHisLeu 827
Db   2614 TCCGTTGCTGGTGGAGGAACATCGAAAGCAGTGTACTTCCCAATTGTGCCGTCTTCAATT 2673
Qy    828 GlyGluIleProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGln 847
Db   2674 GGAGAGATCCCGGTACATATAAGCGCTATTGCATCCCAAGGTGGTGATGCAGTAGAGATG 2733
Qy    848 MetIleLeuValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAsp 867
Db   2734 AACTTGAGAGTTGATCCACAAGGATATAAAGTAGATAGAAATATTCCATTTGTCATTGAC 2793
Qy    868 LeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsn 887
Db   2794 CTCAACAATAACTCTTCGGATTTCTCC---AAAAATCTTGAGCTTATCTGGCCAAACGAC 2850
Qy    888 ThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSer 907
```

```
Db       2851 GTGGTTGATGGATCGCAGAAGGCAAGATTAGATGTGATTGGTGATATGATGGGACCAGTG 2910

Qy        908 IleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIle 927
              :::|||             |||::::::||||||||||||||||||||||||||||:::
Db       2911 TTAAACAATGCTCATAAGCTCGTACAAATGCCATATGGATGTGGAGAACAGAATATGTTG 2970

Qy        928 AsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAsp 947
              |||       |||||||||   ::::::   ||||||             :::
Db       2971 AACTTGGTTCCAAATATTTTGGTGGTCAAGTACTTAAGAGCAACAAATAGAAATGAGAGC 3030

Qy        948 AsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyr 967
                 |||:::    |||||||:::   |||:::    ||||||    |||||||||||    |||
Db       3031 CAATTAGAGACAAAGGCTATCAAGTTTATTGAGCAGGGAATTCAGAGGGAGTTGACGTAC 3090

Qy        968 GlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrp 987
              :::|||   |||   ||||||||||||||||||:::    |||     :::|||||||||||
Db       3091 AAACGTGCTGACAACTCTTTCTCAGCTTTTGGCGATAGTGATAAAGCTGGTTCAACGTGG 3150

Qy        988 LeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGln 1007
              |||:::|||||||||:::|||    |||      |||      |||         :::|||
Db       3151 CTCACCGCATTTGTCGTTCGATCGTTCCATCACGCCAAACAATACGCATTCGTGGATCCT 3210

Qy       1008 AsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPhe 1027
              ||||||:::     |||         :::|||       |||       :::|||    |||
Db       3211 AATGTTATTTCCCGAGCAGTCGCATTCTTGAACTCTCAACAAATGGAGTCTGGTGCATTT 3270

Qy       1028 TrpAspProGlyArgValIleHisSerGluLeuGlnGlyGly---AsnLysSerProVal 1046
              :::    |||    |||    |||     ::::::||||||||                |||
Db       3271 GCCGAACGTGGTGAAGTTCATCACAAGGATATGCAGGGAGGTGCTCAGGACGGGGGTGTA 3330

Qy       1047 ThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIle 1066
                 |||||||||::::::::    |||:::|||              :::       :::
Db       3331 GCTCTAACGGCATTTGTGCTAATCTCGATTTTG---------------GAGAATGGAATG 3375

Qy       1067 AspValGlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsn 1086
              :::      :::::::::    :::||||||                           :::
Db       3376 GAGAATGGAAAGGCGGTGACGTATTTGGAAAAGCATTTGGATGAAGTATCTGGAAATGCG 3435

Qy       1087 TyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGlu 1106
              ||||||:::|||::::::   |||||||||            |||    :::|||    :::
Db       3436 TATACAATGGCTGTAGTGGCATATGCACTTCAATTGGCAAAAAGTAAGCAAGCTGGAAAG 3495

Qy       1107 AlaLeuAsnMetLeuThrTrpArg------AlaGluGlnGluGlyGlyMetGlnPheTrp 1124
              |||          |||                |||:::    |||    ::::::|||
Db       3496 GCATTCGAGAATTTGAAGAAACATAAGATCGTTGAGAAAAGTGGAGATGTGAAGTTTGCC 3555

Qy       1125 ValSerSerGlu------SerLysLeuSerAspSer------------TrpGlnProArg 1138
              :::    :::          ||||||    :::|||              :::|||  |||
Db       3556 AGTGCTCAGAAAAAAGTGGAGAAACTAAAAGAATCGAGAGCGTATATGTTCCAGGCCCGT 3615

Qy       1139 SerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThr 1158
                 :::|||||||||       :::||||||:::||||||:::       |||    ||||||
Db       3616 CCAGTAGACATTGAAACTACCTCTTACGCAGTTCTTTCTTACCTCGCCCAAAATCAAACC 3675

Qy       1159 SerGluGlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPhe 1178
              ||||||    :::    |||:::|||||||||       |||||||||    ||||||||||||
Db       3676 TCAGAATCTCTCTCAATTATTCGTTGGTTAGTATCCCAACGCAACGAGCTAGGAGGATTC 3735

Qy       1179 AlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMet 1198
                 |||||||||||||||    :::||||||:::|||||||||    :::||||||:::
Db       3736 ACTTCCACTCAAGACACTGTCATGGCTCTTCAAGCTCTCTCTTCCTACGCAGCAGTCACT 3795

Qy       1199 AsnThrGluArgThrAsnIleGlnValThrVal---ThrGlyProSerSerProSer--- 1216
                 :::::::::          ||||||||||:::     |||     :::    |||
Db       3796 TATTCCGACAAACACACAAGTCAAGTAACAATTCTCAACGGAAAACATACTCACTCATTT 3855

Qy       1217 ---------------ProLeuAlaValValGlnProMet--------------------- 1224
                                 |||:::|||
Db       3856 GATATCAACATAAGAAACGCAATTGTTCTCCAATCCTATCAACTATCTTCATTGAATGAT 3915

Qy       1225 AlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyr 1244
              ||||||:::|||:::|||||||||    |||           ||||||:::       |||
Db       3916 GCAGTTTCAATTAATGCCAATGGAACTGGTGTGGTCTTCGCTCAGTTGAGTTATTCTTAC 3975

Qy       1245 AsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAlaPhe 1264
                                          |||   |||:::    :::        |||
Db       3976 ---------------------------TACAGGGACTCTTTGAATGACGATGCTCCGTTC 4008

Qy       1265 AspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsnVal 1284
                       :::||||||    :::              |||    :::    |||::::::
Db       4009 TTCTGCTCCCAGGAGATCAAGGAAATAAGAGCTGGA---AATCGACTGCAATTGGATTTG 4065

Qy       1285 CysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeu 1304
              |||    :::::::::    ||||||:::|||    |||||||||    |||::::::   |||
Db       4066 TGTTGCAACTATACCCGTCCTGGAAAGTCTAACATGGCTCTAGCCGAAATTGATGCCTTG 4125

Qy       1305 SerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLys--------- 1321
              ||||||:::          :::|||    :::             ::::::::
Db       4126 AGTGGTTACCGATTCGATGCGGAACAGGTGCATACTTTGACAAGTATTGAGGATTTGCAA 4185

Qy       1322 ------LysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGlu 1339
                    ::::::|||    |||       |||:::||||||:::|||    :::    :::
Db       4186 AGGGTGGAGATGGAAAAGGACGACACTAAGATGAACGTATATTTCAACCCGCTCGGTGGC 4245

Qy       1340 ThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAsp 1359
                       |||:::::::::    :::          ::::::|||::::::    :::
Db       4246 CGTCCGGTATGTCTCTCGTTATACTCTGATGTCACTTATCAAGTTGCCGATCAAAAACCT 4305

Qy       1360 AlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsn 1379
              |||:::       :::|||||||||||||:::|||       |||           :::|||
Db       4306 GCCAACTTCCGATTAGTGGATTACTACGATCCGGAGGAGCAGCTCAAAATGACATACGCC 4365

Qy       1380 SerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgPro 1398
              ::::::          |||    :::    |||          :::    |||    |||
Db       4366 GCCAAGCAAACACGATCGCTGCAAGAGAAATGTGGG------GAAGACTGTTGGCCA 4416
```

RESULT 15
AX054954

```
LOCUS       AX054954                 4560 bp    DNA     linear   PAT 13-JAN-2001
DEFINITION  Sequence 69 from Patent WO0073328.
ACCESSION   AX054954
VERSION     AX054954.1  GI:12228338
KEYWORDS    .
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1
  AUTHORS   van Criekinge,W., Roelens,I., Bogaert,T. and Verwaerde,P.
  TITLE     Unc-5 constructs and screening methods
  JOURNAL   Patent: WO 0073328-A 69 07-DEC-2000;
            Devgen NV (BE)
FEATURES             Location/Qualifiers
     source          1..4560
                     /organism="Caenorhabditis elegans"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:6239"
BASE COUNT     1326 a    955 c   1084 g   1195 t
ORIGIN

Alignment Scores:
Pred. No.:              4.54e-112      Length:         4560
Score:                  1850.00        Matches:        477
Percent Similarity:     50.79%         Conservative:   290
Best Local Similarity:  31.59%         Mismatches:     563
Query Match:            25.18%         Indels:         180
DB:                     6              Gaps:           44

US-10-020-095-4 (1-1428) x AX054954 (1-4560)

Qy         15 ValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThrAlaProGly 34
              |||    |||    |||          |||    |||          :::::::::    ||||||
Db         82 GTGTCAACAACCGCGGCGCCAGTTAAACCAGCC---ACTTATATGCTTGTCGCCCCAGCA 138

Qy         35 IleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHisCysProSer 54
              ::::::||||||                   ::::::    :::    ||||||:::
```

```
Db      139 GTAGTGCGACCCGATCAACCGTTCTCCGTTTGCATGAACCTTCTGAAACAAGCTACGGAT 198

Qy       55 GlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeu 74
            :::      :::            ::::::|||       |||   |||::::::
Db      199 GAAGACATGATTGTTCGGATTGAAGTGAGAACTGAACGGAACGAGACGATTGCTGCTCGA 258

Qy       75 GluAlaGluGlyValPheGluLysGlySerPheLysThrLeuThrLeuProSerLeuPro 94
                  :::   :::   |||      :::|||::::::|||       :::|||
Db      259 GTTATAAGCAATTTG---AAGCCGGGAATCGCACAGACTGTCAGCTTGAGTGAGATGCCT 315

Qy       95 LeuAsnSerAlaAsp------GluIleTyrGluLeuArgValThrGlyArgThrGlnAsp 112
               |||             :::   |||:::|||    :::   |||  |||   :::
Db      316 GCTCAATCACTAACTCCTAGGCAAAGCTATAAGCTCTACATCCGCGGGGAGACTCTTAAC 375

Qy      113 ---GluIleLeuPheSerAsnSerThrArgLeuSerPheGluThrLysArgIleSerVal 131
               |||::::::|||   |||         |||  ::::::   |||   ::::||||||
Db      376 GCAGAACTCATATTTGAAAACGAAAATGAGTTAAAGTACGATCAAAAAGCTTTATCAGTT 435

Qy      132 PheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluValLysPheArgIleVal 151
            |||||||||||||||:::|||:::|||:::|||            |||::::::|||   :::
Db      436 TTCATCCAAACAGATCGAGCAATCTATCGGCCTGCTTCTCTTGTTCGATACCGGGCAATT 495

Qy      152 ThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIleLeuIleLysAspPro 171
               :::   ||||||   |||||||||         :::      |||   |||   ||||||
Db      496 GTAGTGAAATCGGACTTGAAACCTTACGTCGGAAACGCTACTATTAAAATTTTCGATCCA 555

Qy      172 LysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAspLeuGlyValIleSerLys 191
               |||||||||   |||    :::               |||   ||||||   |||
Db      556 AGCAGAAATCTGATTTCACAGACGATTGGTGTAACTTTGGATAGAGGAGTATACAGCGGC 615

Qy      192 ThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSerIleGlnValGln------ 209
               ||||||:::            :::||||||||||||   |||:::|||:::
Db      616 GAGCTCCAACTTGCCGAAGAGACTCTCCTTGGAGATTGGTTATCGAGGTGGAAACCTCG 675

Qy      210 ------ValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrValLeuPro 227
                  |||   |||::::::            ||||||   |||      ||||||||||||
Db      676 AATGGAGTCCAGGACAAGTCG---------TCATTTACGGTAGATACATACGTGCTTCCC 726

Qy      228 LysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSerLysHisLeuAsn 247
            ||||||||||||   ::::::|||         :::   ::::::|||            |||:::
Db      727 AAATTTGAAGTGAATATCAAGACGTCTTCGTTTATAACTATTAACGAT---GATTTGTCA 783

Qy      248 GlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAspValThrLeuThr 267
               :::   ||||||||||||||||||   |||   |||               ::::::
Db      784 GTGTTCGTCGATGCAAAGTACACATATGGAAAAGGCGTTGCTGGAAAGGCCAAAGTATCT 843

Qy      268 Phe---LeuProLeuSerPheTrp-------------------------------Gly 275
                  ||||||      |||                                     |||
Db      844 CTAGAGCTTCCATGGCACAGATGGCATGCGATGGTTCCCACTATCATTGACGAAAATGGA 903

Qy      276 LysLysLysAsn------------IleThrLysThrPheLysIleAsn------GlySer 289
               ||||||             :::   :::|||   |||:::|||       |||
Db      904 GTCAAGAAGGAAGAGGAGCTGATGGTCGAGCGTACTGTAAAGCTGAACAGACAAGGAGAA 963

Qy      290 AlaAsnPheSerPheAsnAspGluGluMetLysAsnValMetAspSerSerAsnGlyLeu 309
            |||         |||:::::::::|||:::|||
Db      964 GCGGCAGTTGTTTTCTCGAATGATGAGCTCAAGAGG------------------------ 999

Qy      310 SerGluTyrLeuAspLeuSerSerProGlyProValGluIleLeuThrThrValThrGlu 329
               :::   |||               |||   :::   |||:::   :::|||||||||
Db     1000 CACAAGTTGCTACATGAATGGGGAGGAGGTTCAATTCGAATTGTTGCCTCAGTCACTGAA 1059

Qy      330 SerValThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyr 349
            :::|||   |||   ||||||         :::         :::         :::
Db     1060 GACATCACTGAAATCGAAAGAAACGCTACTCATCAAATTTCAACGTTCCGTGAGGAGGTC 1119

Qy      350 IleIleGluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThr 369
               ::::::                        ||||||   |||   :::
Db     1120 AAACTTGATGTTGAAAAGCAGGGAGACACCTTCAAGCCAGGACTCACTTATAATGTTGTA 1179

Qy      370 ValLysValThrArgAlaAspGlyAsn------GlnLeuThrLeuGluGluArgArgAsn 387
            |||   :::   :::   |||                 :::   ||||||   :::|||
Db     1180 GTTGCACTGAAACAAATGGATGACACTCCCGTGAAGGCAACTTTGCCAAAGAGA------ 1233

Qy      388 AsnValValIleThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSer 407
                  |||         |||   ||||||            ::::::            |||
Db     1234 ------GTTCAGGTATCAACATTCTACAACTATCCA------TACAATCATGACACTTCA 1281

Qy      408 GlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThr 427
                  |||:::   :::      :::   :::                  :::
Db     1282 TCTCTGCAAGAGGAAAAAGAAACAAAAATTGTGGAAGTTGACGCTCACGGCACCTCGGTT 1341

Qy      428 PheLysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPhe 447
               ::::::   ||||||            |||:::   ::::::            |||
Db     1342 CTCACCCTCCAACCACCAATCAACTGCACTTCTGCTCGCATTGAGGCTCAT---TATGAT 1398

Qy      448 LeuGlySerLysSerSerMetAlaValHisSerLeuPheLys------------------ 461
            :::|||   |||   :::            ::::::
Db     1399 ATTGGTGGCAAGGACAACTTCACAGCCACCCCAATATATTCCTCGCTATATGTCGAGGCA 1458

Qy      462 ------SerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsn---IleLys 478
                  ||||||:::|||:::::::::||||||         :::|||         :::
Db     1459 GCCGTGAGCCCAACGAAGAGTTTCCTGCAGCTTTTGGCTGATAATGAGGGAGCTGTTGAT 1518

Qy      479 ValGlySerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyr 498
            ||||||               :::      :::   |||      ::::::|||
Db     1519 GTAGGAAAGAGTCTTAGCTTCTCGTTGAAGGCTACCCAGCCGCTGTCGACAATTACCTAT 1578

Qy      499 MetValValSerArgGlyGlnLeuVal---------------AlaValGlyLysGlnAsn 513
               |||:::||||||      :::|||                  |||      ::::::
Db     1579 CAGGTCATGTCTCGTTCGAACATTGTGGTGTCTCAACAAATGACTGTTAATTCCGAACAT 1638

Qy      514 SerThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleVal 533
            :::|||:::   |||               :::      ||||||:::      :::||||||
Db     1639 GCGACCATT---AGCTTCCCAGCAACGGCTAACATGGCCCCAAAGTCACGGTTGATCGTC 1695

Qy      534 Tyr---TyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGln 552
            |||      ||||||      |||::::::   |||   |||                  |||:::
Db     1696 TATGCCATTATTGAGAGTAGTCAAGAAGTTCTTGTTGATGCACTTGACTTTAAAGTAGAA 1755

Qy      553 LeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGluProSerGlu 572
               :::|||:::|||::::::   |||         |||            ||||||   :::
Db     1756 GGAATATTCCAAAACCAAGTGGCACTATCAATTGACAAACAAGCAGTCGAGCCAGGACAG 1815

Qy      573 LysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaVal 592
               |||   :::::::::   :::      :::|||   |||||||:::::   |||
Db     1816 AATGTGAAGTTTAAGGTCACA---TCTGACAAGAACTCTTTTGTTGGGCTTCTTGTGGTA 1872

Qy      593 AspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValValHis 612
            |||:::|||||||   |||:::            |||||||||||   |||   |||
Db     1873 GATCAGAGTGTTCTCTTGTTAAAAACTGGCAATGATATCACTAGAGAAAAAGTTGAGCAA 1932

Qy      613 GluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPhe------------------ 626
            :::||||||   |||::::::            |||   |||
Db     1933 GACTTGGAGAATTATGACTCTAATAATGTTGGAGGAGGATTCGGGGGTCCACGGCCATGG 1992

Qy      627 -------------------------------------------------------Met 627

Db     1993 GAGGCAATTGATAGAAAGAAGAGGTCCATTTGGAGACCTTGGTGGGGAATTGGAGGGAGT 2052

Qy      628 AsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeuThrAspAla--------- 644
            ::::::   ::::::|||            ||||||   ||||||||||||||
Db     2053 GATGCTCAATCAATCTTTTCTAATGCTGGGCTTGTAGTTTTGACCGACGCCCCTCTTGTAC 2112

Qy      645 ------------------------------------------------AsnLeuThr 647
                                                               ||||||
Db     2113 CGCGAGCCTCAACGTGAATTTATGTCAGAAAGAAGACTTAATACACCCGGAGGTCTTACA 2172

Qy      648 LysAspTyrIleAspGlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGlu 667
               :::||||||            ||||||   |||         |||
Db     2173 GTAATGATGATGGATGGTGCTCCAGGAATGGCTGAAGCCGCA------TTCGCCGCTCCG 2226

Qy      668 AsnGluGlyHisIleValAspIleHisAspPheSerLeuGlySerSer---------Pro 684
               |||                              |||||||||               |||
Db     2227 CCAATGGGC----------------------------GGGTCGTCGCCTCCACCGCCG 2256
```

```
Qy    685 HisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArg 704
Db   2257 ACAGTTCGAAAGTTCTTCCCACACACCTGGATTTGGTCGGATCTCAAC---TCTACTAGT 2313
Qy    705 IleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGly 724
Db   2314 GGAGAAGTTGAGATGGAAATTGAAGCTCCGGACACCATCACCTCGTGGGTCGCCTCCACT 2373
Qy    725 PheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAla 744
Db   2374 TTTGCAATCAATGAGGAAAATGGTCTCGGTGTGGCTCCAACGACTTCAAAACTTCGCGTG 2433
Qy    745 PheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPhe 764
Db   2434 TTTCGTCCATTTTTTATTCAACTCAATCTTCCATATGCTGTCCGTCGTGGAGAAAAGTTT 2493
Qy    765 AlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIle 784
Db   2494 GCACTTTTGGTGCTTGTTTTCAATTATATGGAAAAGGAGCAAGATGTAACAGTGACTCTG 2553
Qy    785 ------GluLysSerAspLysPheAspIleLeuMetThrSerSerGluIle--------- 799
Db   2554 AAGTATGATAAAGACTCTGGTTACGATCTCCTGAAGAAAGATGGAACTGTAGTAAGGCGT 2613
Qy    800 AsnAlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThr------ 817
Db   2614 GACGAGGTTGGCCAACAAAATGTTCGTATTGTGTCCGTTGCTGGTGGAGGAACATCGAAA 2673
Qy    818 ---ValLeuPheProIleArgProThrHisLeuGlyGluIleProIleThrValThrAla 836
Db   2674 GCAGTGTACTTCCCAATTGTGCCGTCTTCAATTGGAGAGATCCCGGTACATATAAGCGCT 2733
Qy    837 LeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
Db   2734 ATTGCATCCCAAGGTGGTGATGCAGTAGAGATGAACTTGAGAGTTGATCCACAAGGATAT 2793
Qy    857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
Db   2794 AAAGTAGATAGAAATATTCCATTTGTCATTGACCTCAACAATAACTCTTCGGATTTCTCC 2853
Qy    877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
Db   2854 ---AAAAATCTTGAGCTTATCTGGCCAAACGACGTGGTTGATGGATCGCAGAAGGCAAGA 2910
Qy    897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
Db   2911 TTAGATGTGATTGGTGATATGATGGGACCAGTGTTAAACAATGCTCATAAGCTCGTACAA 2970
Qy    917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
Db   2971 ATGCCATATGGATGTGGAGAACAGAATATGTTGAACTTGGTTCCAAATATTTTGGTGGTC 3030
Qy    937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
Db   3031 AAGTACTTAAGAGCAACAAATAGAAATGAGAGCCAATTAGAGACAAAGGCTATCAAGTTT 3090
Qy    957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
Db   3091 ATTGAGCAGGGAATTCAGAGGGAGTTGACGTACAAACGTGCTGACAACTCTTTCTCAGCT 3150
Qy    977 PheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPhe 996
Db   3151 TTTGGCGATAGTGATAAAGCTGGTTCAACGTGGCTCACCGCATTTGTCGTTCGATCGTTC 3210
Qy    997 LeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrp 1016
Db   3211 CATCACGCCAAACAATACGCATTCGTGGATCCTAATGTTATTTCCCGAGCAGTCGCATTC 3270
Qy   1017 LeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSer 1036
Db   3271 TTGAACTCTCAACAAATGGAGTCTGGTGCATTTGCCGAACGTGGTGAAGTTCATCACAAG 3330
Qy   1037 GluLeuGlnGlyGly---AsnLysSerProValThrLeuThrAlaTyrIleValThrSer 1055
Db   3331 GATATGCAGGGAGGTGCTCAGGACGGGGGTGTAGCTCTAACGGCATTTGTGCTAATCTCG 3390
Qy   1056 LeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeu 1075
Db   3391 ATTTTG---------------GAGAATGGAATGGAGAATGGAAAGGCGGTGACGTATTTG 3435
Qy   1076 GluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAla 1095
Db   3436 GAAAAGCATTTGGATGAAGTATCTGGAAATGCGTATACAATGGCTGTAGTGGCATATGCA 3495
Qy   1096 LeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArg--- 1114
Db   3496 CTTCAATTGGCAAAAAGTAAGCAAGCTGGAAAGGCATTCGAGAATTTGAAGAAACATAAG 3555
Qy   1115 ---AlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGlu------SerLysLeu 1131
Db   3556 ATCGTTGAGAAAAGTGGAGATGTGAAGTTTGCCAGTGCTCAGAAAAAAGTGGAGAAACTA 3615
Qy   1132 SerAspSer------------TrpGlnProArgSerLeuAspIleGluValAlaAlaTyr 1147
Db   3616 AAAGAATCGAGAGCGTATATGTTCCAGGCCCGTCCAGTAGACATTGAAACTACCTCTTAC 3675
Qy   1148 AlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrp 1167
Db   3676 GCAGTTCTTTCTTACCTCGCCCAAAATCAAACCTCAGAATCTCTCTCAATTATTCGTTGG 3735
Qy   1168 LeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAla 1187
Db   3736 TTAGTATCCCAACGCAACGAGCTAGGAGGATTCACTTCCACTCAAGACACTGTCATGGCT 3795
Qy   1188 LeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnVal 1207
Db   3796 CTTCAAGCTCTCTCTTCCTACGCAGCAGTCACTTATTCCGACAAACACACAAGTCAAGTA 3855
Qy   1208 ThrVal---ThrGlyProSerSerProSer------------------ProLeuAlaVal 1220
Db   3856 ACAATTCTCAACGGAAAACATACTCACTCATTTGATATCAACATAAGAAACGCAATTGTT 3915
Qy   1221 ValGlnProMet--------------------AlaValAsnIleSerAlaAsnGlyPhe 1233
Db   3916 CTCCAATCCTATCAACTATCTTCATTGAATGATGCAGTTTCAATTAATGCCAATGGAACT 3975
Qy   1234 GlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArg 1253
Db   3976 GGTGTGGTCTTCGCTCAGTTGAGTTATTCTTAC-------------------------- 4008
Qy   1254 ArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsn 1273
Db   4009 TACAGGGACTCTTTGAATGACGATGCTCCGTTCTTCTGCTCCCAGGAGATCAAGGAAATA 4068
Qy   1274 LysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArg 1293
Db   4069 AGAGCTGGA---AATCGACTGCAATTGGATTTGTGTTGCAACTATACCCGTCCTGGAAAG 4125
Qy   1294 SerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAla 1313
Db   4126 TCTAACATGGCTCTAGCCGAAATTGATGCCTTGAGTGGTTACCGATTCGATGCGGAACAG 4185
Qy   1314 IleSerLeuSerGluThrValLys---------------LysValGluTyrAspHisGly 1328
Db   4186 GTGCATACTTTGACAAGTATTGAGGATTTGCAAAGGGTGGAGATGGAAAAGGACGACACT 4245
Qy   1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
Db   4246 AAGATGAACGTATATTTCAACCCGCTCGGTGGCCGTCCGGTATGTCTCTCGTTATACTCT 4305
Qy   1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
Db   4306 GATGTCACTTATCAAGTTGCCGATCAAAAACCTGCCAACTTCCGATTAGTGGATTACTAC 4365
Qy   1369 GluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCysAsp 1388
```

```
Db       4366 GATCCGGAGGAGCAGCTCAAAATGACATACGCCGCCAAGCAAACACGATCGCTGCAAGAG 4425

Qy       1389 LeuCysSerAspValGlnGlyCysArgPro 1398
                 |||         :::   |||    |||
Db       4426 AAATGTGGG------GAAGACTGTTGGCCA 4449

Search completed: January 17, 2004, 08:32:16
Job time : 11145 secs
```

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         January 16, 2004, 20:57:28 ; Search time 753 Seconds
                                             (without alignments)
                                             5119.253 Million cell updates/sec

Title:          US-10-020-095-4
Perfect score:  7348
Sequence:       1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters:     5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10020095/runat_15012004_163821_18701/app_query.fasta_1.1607
-DB=N_Geneseq_19Jun03 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10020095_@CGN_1_1_603_@runat_15012004_163821_18701 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      N_Geneseq_19Jun03:*
                1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
                25:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	7331	99.8	4369	25	AAD49434	Human r150 gene #1
	2	7326.5	99.7	4335	24	AAL49815	Human platelet all
	3	7326.5	99.7	4761	24	ABQ79964	Human CD109 K1 pro
	4	7326.5	99.7	5882	25	AAD49440	Human blood cell s
	5	7326.5	99.7	5895	24	ABQ79966	Human CD109 K1-H7
	6	7318.5	99.6	5883	25	AAD49435	Human r150 DNA #1.
	7	7317.5	99.6	4335	24	AAL49816	Human platelet all
	8	7317.5	99.6	4761	24	ABQ79965	Human CD109 K1 var
	9	7317.5	99.6	5895	24	ABQ79967	Human CD109 K1-H7
	10	7307.5	99.4	4335	21	AAA60199	Hydrophobic domain
	11	7307.5	99.4	4473	21	AAA62010	Hydrophobic domain
	12	7081	96.4	4146	25	AAD49436	Human blood cell s
	13	7068.5	96.2	4197	25	AAD49437	Human r150 DNA #2.
	14	5470	74.4	3535	24	ABQ79968	Human CD109 K15 pr
	15	5461	74.3	3535	24	ABQ79969	Human CD109 K15 va
	16	1854.5	25.2	4527	22	AAC90951	C. elegans alpha-2
	17	1850	25.2	4560	22	AAC90952	C. elegans alpha-2
	18	1819.5	24.8	1448	24	ABL90478	Human polynucleoti
c	19	1767.5	24.1	1459	22	ABA08825	Human secreted pro
	20	1533	20.9	5816	21	AAC74484	Human ORFX ORF39 p
	21	1461.5	19.9	2403	24	ABV77330	Alpha-1 proteinase
	22	1441	19.6	4577	21	AAA52113	Human alpha-2-macr
	23	1440	19.6	4425	22	AAC90944	Human alpha-2 macr
	24	1440	19.6	4577	22	AAD07359	Human DNA encoding
	25	1440	19.6	4577	24	AAD45346	Human anti-microbi
	26	1440	19.6	4577	24	ABN95738	Gene #2236 used to
	27	1440	19.6	4577	24	ABK64513	Human benign prost
	28	1440	19.6	4577	24	ABK12952	DNA encoding human
	29	1440	19.6	4577	24	ABK24095	Human alpha2 macro
	30	1440	19.6	4578	25	ABX74411	Human cDNA sequenc
	31	1440	19.6	4578	25	ABX74412	Human cDNA sequenc
	32	1440	19.6	4809	24	AAS94923	Human DNA sequence
	33	1438	19.6	4527	22	AAH74529	A human alpha-2 ma
	34	1438	19.6	4527	22	AAH74530	A human alpha-2 ma
	35	1430	19.5	3612	23	ABL19705	Drosophila melanog
	36	1428.5	19.4	4569	12	AAQ11176	Clone encoding rec
	37	1425	19.4	4599	12	AAQ11581	Sequence encoding
	38	1422	19.4	4488	24	ABS59327	Human alpha-2-macr
	39	1412.5	19.2	4595	24	ABK63624	Rat sequence diffe
	40	1406	19.1	4278	23	ABL30183	Drosophila melanog
	41	1401	19.1	4487	24	ABK92035	DNA encoding novel
	42	1400	19.1	4615	23	AAS92193	DNA encoding novel
	43	1400	19.1	4615	24	ABN97232	Gene #3730 used to
	44	1376.5	18.7	4660	24	ABQ93897	Human pregnancy zo
c	45	1376.5	18.7	6324	23	ABL19704	Drosophila melanog

ALIGNMENTS

```
RESULT 1
AAD49434
ID   AAD49434 standard; DNA; 4369 BP.
XX
AC   AAD49434;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 gene #1.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
```

```
FT   CDS             26..4312
FT                   /*tag=  a
FT                   /product= "Human r1520 protein #1"
FT                   /transl_except= (pos:2132..2134, aa:Xaa)
FT                   /note= "Xaa corresponds to Ser, Tyr"
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   P-PSDB; AAE32012.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 91-93; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is a gene encoding
CC   human r150 protein.
XX
SQ   Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;

Alignment Scores:
Pred. No.:              0          Length:        4369
Score:                  7331.00    Matches:       1425
Percent Similarity:     99.86%     Conservative:  1
Best Local Similarity:  99.79%     Mismatches:    2
Query Match:            99.77%     Indels:        0
DB:                     25         Gaps:          0

US-10-020-095-4 (1-1428) x AAD49434 (1-4369)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 85

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 145

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 205

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 265

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 325

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 385

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 445

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 505

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 565

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 625

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAG 685

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 745

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 805

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 865

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 925

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 985

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1045

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1105

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1165

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1225

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1285

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1345

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1405

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1406 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1465
```

```
Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1466 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1525

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1526 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1585

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1586 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1645

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1646 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1705

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1706 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1765

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1766 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1825

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1826 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1885

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1886 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1945

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1946 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2005

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2006 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2065

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2066 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2125

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              |||||    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2126 ATGGGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2185

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2186 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2245

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2246 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2305

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2306 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2365

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db       2366 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2425

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2426 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2485

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2486 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2545

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2546 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2605

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2606 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2665

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2666 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2725

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2726 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2785

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2786 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2845

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2846 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2905

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2906 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2965

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2966 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3025

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3026 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3085

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3086 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3145

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3146 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3205

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3206 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3265

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3266 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3325

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3326 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3385

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3386 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3445

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3446 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3505

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3506 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3565

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3566 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3625

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerProLeuAlaVal 1220
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3626 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTG 3685

Qy     1221 ValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeu 1240
            |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db     3686 GTACAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTC 3745

Qy     1241 AsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsn 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3746 AATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAAT 3805

Qy     1261 GlnGluAlaPheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisVal 1280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3806 CAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTG 3865

Qy     1281 AspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlu 1300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3866 GATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAA 3925

Qy     1301 ValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrVal 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3926 GTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTG 3985

Qy     1321 LysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThr 1340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3986 AAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACC 4045

Qy     1341 GlnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAla 1360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4046 CAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCT 4105

Qy     1361 SerValSerIleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSer 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4106 TCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCT 4165

Qy     1381 GluValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGlu 1400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4166 GAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAG 4225

Qy     1401 AspGlyAlaSerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeu 1420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4226 GATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTT 4285

Qy     1421 LeuTyrPheMetGluLeuTrpLeu 1428
            ||||||||||||||||||||||||
Db     4286 CTGTACTTTATGGAACTTTGGCTG 4309

RESULT 2
AAL49815
ID   AAL49815 standard; cDNA; 4335 BP.
XX
AC   AAL49815;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human platelet alloantigen Gova coding sequence.
XX
KW   Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;
KW   SNP; biallelic; bleeding disorder; post-transfusion purpura;
KW   post-transfusion platelet refractoriness; haemostatic; vaccine;
KW   neonatal alloimmune thrombocytopenia; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key                Location/Qualifiers
FT   CDS                1..4335
FT                      /*tag=  a
FT                      /product= "Gova"
FT                      /partial
FT                      /note= "no stop codon"
FT   variation          replace(2108,C)
FT                      /*tag=  b
XX
PN   WO200270738-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00291.
XX
PR   07-MAR-2001; 2001US-273941P.
XX
PA   (SCHU/) SCHUH A.
XX
PI   Schuh A,  Ouwehand W;
XX
DR   WPI; 2002-713460/77.
DR   P-PSDB; AAO19372.
XX
PT   New isolated oligonucleotide binding to a region of CD109 nucleic acid
PT   having a single nucleotide polymorphism that distinguishes a Gova
PT   and/or Govb allele, useful for treating blood disorders e.g. alloimmune
PT   thrombocytopenia  -
XX
PS   Claim 6; Page 23-29; 69pp; English.
XX
CC   The present invention relates to a sequence capable off binding
CC   specifically to a CD109 nucleic acid which has a single nucleotide
CC   polymorphism that distinguishes the Gova and Govb alleles. Detection of
CC   the Gov genotype is useful for detecting whether the subject has or is at
CC   risk of a blood disease, disorder or abnormal physical state, such as
CC   bleeding, or increased risk of bleeding, due to alloimmune destruction of
CC   blood platelets, e.g., post-transfusion purpura, post-transfusion
CC   platelet refractoriness or neonatal alloimmune thrombocytopenia. The
CC   nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
CC   individuals. The present sequence is the Gova coding sequence.
XX
SQ   Sequence 4335 BP; 1296 A; 837 C; 922 G; 1280 T; 0 other;

Alignment Scores:
Pred. No.:              0              Length:         4335
Score:                  7326.50        Matches:        1427
Percent Similarity:     98.82%         Conservative:   1
Best Local Similarity:  98.75%         Mismatches:     0
Query Match:            99.71%         Indels:         17
DB:                     24             Gaps:           1

US-10-020-095-4 (1-1428) x AAL49815 (1-4335)

Qy        1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy       21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy       41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy       61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy       81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy      101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy      121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
```

```
Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAAGCCAGTG 780

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy       1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                        ||||||||||||||||||
Db       3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy       1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4321 ATGGAACTTTGGCTG 4335
```

RESULT 3
ABQ79964

```
ID   ABQ79964 standard; cDNA; 4761 BP.
XX
AC   ABQ79964;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1 protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
```

```
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82165.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 1a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1 cDNA sequence.
XX
SQ   Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

Alignment Scores:
Pred. No.:              0           Length:         4761
Score:                  7326.50     Matches:        1427
Percent Similarity:     98.82%      Conservative:   1
Best Local Similarity:  98.75%      Mismatches:     0
Query Match:            99.71%      Indels:         17
DB:                     24          Gaps:           1

US-10-020-095-4 (1-1428) x ABQ79964 (1-4761)

Qy        1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy       21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy       41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy       61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy       81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy      101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy      121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy      141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy      161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy      181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy      201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy      221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy      241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy      261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy      281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy      301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy      321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy      341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy      361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy      381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy      401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy      421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy      441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy      461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552
```

```
Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                          ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4433 ATGGAACTTTGGCTG 4447
```

```
RESULT 4
AAD49440
ID   AAD49440 standard; DNA; 5882 BP.
XX
AC   AAD49440;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human blood cell surface antigen, CD109 encoding DNA #2.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; blood cell syrface antigen; CD109; ds.
XX
OS   Homo sapiens.
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Disclosure; Fig 17; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human blood cell
CC   surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.
CC   Note: This sequence is stated to encode human CD109 protein shown in
CC   figure 17 of the specification, but this does not appear to be the case.
XX
SQ   Sequence 5882 BP; 1696 A; 1173 C; 1241 G; 1772 T; 0 other;

Alignment Scores:
Pred. No.:              0            Length:        5882
Score:                  7326.50      Matches:       1427
Percent Similarity:     98.82%       Conservative:  1
Best Local Similarity:  98.75%       Mismatches:    0
Query Match:            99.71%       Indels:        17
DB:                     25           Gaps:          1

US-10-020-095-4 (1-1428) x AAD49440 (1-5882)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy       181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy       201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy       221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy       241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy       261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy       281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy       301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy       321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
             ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db      2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy       801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy       821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy       841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy       861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy       881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812
```

```
Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4433 ATGGAACTTTGGCTG 4447
```

```
RESULT 5
ABQ79966
ID   ABQ79966 standard; cDNA; 5895 BP.
XX
AC   ABQ79966;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1-H7 protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1-H7; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1-H7"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
```

```
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82167.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 2a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1-H7 cDNA sequence.
XX
SQ   Sequence 5895 BP; 1709 A; 1173 C; 1241 G; 1772 T; 0 other;

Alignment Scores:
Pred. No.:               0           Length:        5895
Score:                   7326.50     Matches:       1427
Percent Similarity:      98.82%      Conservative:  1
Best Local Similarity:   98.75%      Mismatches:    0
Query Match:             99.71%      Indels:        17
DB:                      24          Gaps:          1

US-10-020-095-4 (1-1428) x ABQ79966 (1-5895)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTG 892

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy      521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy      541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy      561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy      601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy      621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy      641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy      661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy      681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy      701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy      721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy      741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy      761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy      781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
            ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db     2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy      801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy      821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy      841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy      861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy      881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy      901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy      921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy      941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy      961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy      981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy     1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy     1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy     1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy     1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy     1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy     1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy     1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy     1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
            .|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy     1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy     1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy     1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy     1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                      ||||||||||||||||||
Db     3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832
```

```
Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4433 ATGGAACTTTGGCTG 4447
```

```
RESULT 6
AAD49435
ID   AAD49435 standard; DNA; 5883 BP.
XX
AC   AAD49435;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 DNA #1.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "Human r1520 protein #2"
FT                   /transl_except= (pos:2219..2221, aa:Xaa)
FT                   /note= "Xaa corresponds to Ser, Tyr"
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   P-PSDB; AAE32013.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 100-103; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human r150 protein
CC   encoding DNA.
XX
SQ   Sequence 5883 BP; 1696 A; 1173 C; 1241 G; 1772 T; 1 other;

Alignment Scores:
Pred. No.:                0            Length:        5883
Score:                    7318.50      Matches:       1426
Percent Similarity:       98.75%       Conservative:  1
Best Local Similarity:    98.69%       Mismatches:    1
Query Match:              99.60%       Indels:        17
DB:                       25           Gaps:          1

US-10-020-095-4 (1-1428) x AAD49435 (1-5883)
```

```
Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy     181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy     201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy     221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy     241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy     261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy     281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy     301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy     321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy     341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy     361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy     381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy     401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy     421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy     441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy     461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy     481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy     501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy     521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy     541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy     561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy     581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy     601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy     621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy     641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy     661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy     681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy     701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
           |||||    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2213 ATGGGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy     721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy     741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy     761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy     781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
           ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db    2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy     801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy     821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy     841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy     861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy     881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812
```

```
Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4433 ATGGAACTTTGGCTG 4447
```

```
RESULT 7
AAL49816
ID   AAL49816 standard; cDNA; 4335 BP.
XX
AC   AAL49816;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human platelet alloantigen Govb coding sequence.
XX
KW   Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;
KW   SNP; biallelic; bleeding disorder; post-transfusion purpura;
KW   post-transfusion platelet refractoriness; haemostatic; vaccine;
KW   neonatal alloimmune thrombocytopenia; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4335
FT                   /*tag=  a
FT                   /product= "Govb"
FT                   /partial
FT                   /note= "no stop codon"
FT   variation       replace(2108,A)
FT                   /*tag=  b
XX
PN   WO200270738-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00291.
XX
PR   07-MAR-2001; 2001US-273941P.
XX
```

```
PA   (SCHU/) SCHUH A.
XX
PI   Schuh A,  Ouwehand W;
XX
DR   WPI; 2002-713460/77.
DR   P-PSDB; AAO19373.
XX
PT   New isolated oligonucleotide binding to a region of CD109 nucleic acid
PT   having a single nucleotide polymorphism that distinguishes a Gova
PT   and/or Govb allele, useful for treating blood disorders e.g. alloimmune
PT   thrombocytopenia  -
XX
PS   Claim 7; Page 35-41; 69pp; English.
XX
CC   The present invention relates to a sequence capable off binding
CC   specifically to a CD109 nucleic acid which has a single nucleotide
CC   polymorphism that distinguishes the Gova and Govb alleles. Detection of
CC   the Gov genotype is useful for detecting whether the subject has or is at
CC   risk of a blood disease, disorder or abnormal physical state, such as
CC   bleeding, or increased risk of bleeding, due to alloimmune destruction of
CC   blood platelets, e.g., post-transfusion purpura, post-transfusion
CC   platelet refractoriness or neonatal alloimmune thrombocytopenia. The
CC   nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
CC   individuals. The present sequence is the Govb coding sequence.
XX
SQ   Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other;

Alignment Scores:
Pred. No.:              0           Length:        4335
Score:                  7317.50     Matches:       1426
Percent Similarity:     98.75%      Conservative:  1
Best Local Similarity:  98.69%      Mismatches:    1
Query Match:            99.58%      Indels:        17
DB:                     24          Gaps:          1

US-10-020-095-4 (1-1428) x AAL49816 (1-4335)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTAAAAATAAGATAAAG 1680

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy       1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                        ||||||||||||||||||
Db       3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840
```

```
Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4321 ATGGAACTTTGGCTG 4335
```

```
RESULT 8
ABQ79965
ID   ABQ79965 standard; cDNA; 4761 BP.
XX
AC   ABQ79965;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1 variant protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1; variant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1 variant"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82166.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 1b; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1 variant cDNA sequence.
XX
SQ   Sequence 4761 BP; 1419 A; 913 C; 1016 G; 1413 T; 0 other;

Alignment Scores:
Pred. No.:               0           Length:         4761
Score:                   7317.50     Matches:        1426
Percent Similarity:      98.75%      Conservative:   1
Best Local Similarity:   98.69%      Mismatches:     1
Query Match:             99.58%      Indels:         17
DB:                      24          Gaps:           1

US-10-020-095-4 (1-1428) x ABQ79965 (1-4761)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592
```

```
Qy         161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy         181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy         201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy         221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy         241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy         261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy         281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy         301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy         321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy         341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy         361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy         381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy         401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy         421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy         441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy         461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy         481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy         501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy         521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy         541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy         561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy         581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy         601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy         621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy         641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy         661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy         681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy         701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
               ||||||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db        2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy         721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy         741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy         761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy         781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
               ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy         801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy         821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy         841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy         861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy         881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy          901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy          921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy          941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy          961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy          981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy         1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy         1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy         1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy         1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy         1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy         1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy         1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy         1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy         1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy         1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy         1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy         1218 ------------------------------------------LeuAlaValValGlnPro 1223
                                                          ||||||||||||||||||
Db         3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy         1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy         1244 TyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy         1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy         1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy         1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy         1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy         1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy         1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy         1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy         1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy         1424 MetGluLeuTrpLeu 1428
                |||||||||||||||
Db         4433 ATGGAACTTTGGCTG 4447
```

```
RESULT 9
ABQ79967
ID   ABQ79967 standard; cDNA; 5895 BP.
XX
AC   ABQ79967;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1-H7 variant protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1-H7; variant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1-H7 variant"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
```

```
XX
PI     Schuh A,  Sutherland RD;
XX
DR     WPI; 2002-713450/77.
DR     P-PSDB; ABB82168.
XX
PT     New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT     particularly for treating strokes, myocardial infarctions, thrombosis,
PT     thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT     transplantation -
XX
PS     Claim 1; Fig 2b; 156pp; English.
XX
CC     The invention relates to isolated nucleic acid molecules encoding CD109
CC     polypeptides. These nucleic acid molecules include the human cDNA
CC     sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC     CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC     family of thioester-containing proteins.The CD109 polypeptides can be
CC     expressed by standard recombinant methodology. The CD109 nucleic acid,
CC     CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC     for treating or detecting a disease or disorder, e.g. conditions
CC     associated with endothelial activation, platelet activation, activation
CC     of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC     and of the complement system, quantitative or qualitative abnormalities
CC     of platelet function, increased or impaired platelet aggregation and
CC     activation, increased or impaired activation of the coagulation and/or
CC     fibrinolytic systems, or impaired or increased immune activation. These
CC     are also useful for treating cardiovascular disorders, stroke, myocardial
CC     infarction, thrombosis, embolism, peripheral vascular disease,
CC     thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC     transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC     or their antisense nucleotide sequence are useful in gene therapy for
CC     treating these conditions. The present sequence represents the human
CC     CD109 K1-H7 variant cDNA sequence.
XX
SQ     Sequence 5895 BP; 1708 A; 1174 C; 1241 G; 1772 T; 0 other;

Alignment Scores:
Pred. No.:              0          Length:        5895
Score:                  7317.50    Matches:       1426
Percent Similarity:     98.75%     Conservative:  1
Best Local Similarity:  98.69%     Mismatches:    1
Query Match:            99.58%     Indels:        17
DB:                     24         Gaps:          1

US-10-020-095-4 (1-1428) x ABQ79967 (1-5895)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612
```

```
Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              |||||    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       1218 ----------------------------------------LeuAlaValValGlnPro 1223
```

```
                                                      |||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4433 ATGGAACTTTGGCTG 4447

RESULT 10
AAA60199
ID   AAA60199 standard; DNA; 4335 BP.
XX
AC   AAA60199;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.
XX
KW   Human; secreted protein; membrane protein; hydrophobic domain;
KW   proliferation control; differentiation induction; material transport;
KW   biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
KW   immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
KW   haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
KW   autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200029448-A2.
XX
PD   25-MAY-2000.
XX
PF   17-NOV-1999;   99WO-JP06412.
XX
PR   17-NOV-1998;   98JP-0326255.
PR   22-DEC-1998;   98JP-0364315.
PR   16-MAR-1999;   99JP-0069811.
PR   27-APR-1999;   99JP-0119299.
PR   19-MAY-1999;   99JP-0138169.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Kimura T;
XX
DR   WPI; 2000-387753/33.
DR   P-PSDB; AAB12127.
XX
PT   Proteins comprising hydrophobic regions, such as secretory and membrane
PT   proteins, useful in research and diagnostics and having various
PT   activities e.g. immunomodulatory, antiinflammatory, chemokinetic,
PT   hemostatic, thrombolytic -
XX
PS   Claim 3; Page 244-246; 410pp; English.
XX
CC   Secretory proteins play important roles in the proliferation control, the
CC   differentiation induction, the material transport and the biophylaxis of
CC   cells. Membrane proteins have important roles as signal receptors, ion
CC   channels and transporters. The present sequence is the coding sequence
CC   for a human protein which has at least one hydrophobic domain. The
CC   protein encoded by the present sequence may be a secretory or a membrane
CC   protein. The encoded protein may have cytokine and cell
CC   proliferation/differentiation activity, immune stimulating or suppressing
CC   activity, haematopoiesis activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, anti-inflammatory activity and tumour
CC   inhibition activity. The present sequence could therefore be used for
CC   treatment of autoimmune disease, Alzheimer's disease, Parkinson's
CC   disease, and cancer via gene therapy.
XX
SQ   Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other;

Alignment Scores:
Pred. No.:              0           Length:        4335
Score:                  7307.50     Matches:       1424
Percent Similarity:     98.62%      Conservative:  1
Best Local Similarity:  98.55%      Mismatches:    3
Query Match:            99.45%      Indels:        17
DB:                     21          Gaps:          1

US-10-020-095-4 (1-1428) x AAA60199 (1-4335)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              |||||||||||||||||||||||||||||||||||||||||||||     ||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTTCCCTGGACCA 960

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
              |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
              ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GCTTCTGATGCTATCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
```

```
Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy       1218 ---------------------------------------LeuAlaValValGlnPro 1223
                                                     ||||||||||||||||||
Db       3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4321 ATGGAACTTTGGCTG 4335

RESULT 11
AAA62010
ID   AAA62010 standard; DNA; 4473 BP.
XX
AC   AAA62010;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.
XX
KW   Human; secreted protein; membrane protein; hydrophobic domain;
KW   proliferation control; differentiation induction; material transport;
KW   biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
KW   immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
KW   haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
KW   autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200029448-A2.
XX
PD   25-MAY-2000.
XX
PF   17-NOV-1999;   99WO-JP06412.
```

```
XX
PR   17-NOV-1998;   98JP-0326255.
PR   22-DEC-1998;   98JP-0364315.
PR   16-MAR-1999;   99JP-0069811.
PR   27-APR-1999;   99JP-0119299.
PR   19-MAY-1999;   99JP-0138169.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Kimura T;
XX
DR   WPI; 2000-387753/33.
DR   P-PSDB; AAB12127.
XX
PT   Proteins comprising hydrophobic regions, such as secretory and membrane
PT   proteins, useful in research and diagnostics and having various
PT   activities e.g. immunomodulatory, antiinflammatory, chemokinetic,
PT   hemostatic, thrombolytic -
XX
PS   Claim 4; Page 253-261; 410pp; English.
XX
CC   Secretory proteins play important roles in the proliferation control, the
CC   differentiation induction, the material transport and the biophylaxis of
CC   cells. Membrane proteins have important roles as signal receptors, ion
CC   channels and transporters. The present sequence is the coding sequence
CC   for a human protein which has at least one hydrophobic domain. The
CC   protein encoded by the present sequence may be a secretory or a membrane
CC   protein. The encoded protein may have cytokine and cell
CC   proliferation/differentiation activity, immune stimulating or suppressing
CC   activity, haematopoiesis activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, anti-inflammatory activity and tumour
CC   inhibition activity. The present sequence could therefore be used for
CC   treatment of autoimmune disease, Alzheimer's disease, Parkinson's
CC   disease, and cancer via gene therapy.
XX
SQ   Sequence 4473 BP; 1328 A; 867 C; 952 G; 1326 T; 0 other;

Alignment Scores:
Pred. No.:              0          Length:        4473
Score:                  7307.50    Matches:       1424
Percent Similarity:     98.62%     Conservative:  1
Best Local Similarity:  98.55%     Mismatches:    3
Query Match:            99.45%     Indels:        17
DB:                     21         Gaps:          1

US-10-020-095-4 (1-1428) x AAA62010 (1-4473)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         45 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 104

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 164

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 224

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 284

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 344

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 404

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 464

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 524

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 584

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        585 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 644

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 704

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        705 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 764

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        765 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 824

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        825 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTGGGGAAAGAAGAAAAATATT 884

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        885 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 944

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||    ||||||||||
Db        945 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTTCCCTGGACCA 1004

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1005 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1064

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1065 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1124

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1125 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1184

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1185 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1244

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1245 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1304

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1305 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1364

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1365 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1424

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1425 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1484

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1485 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1544

Qy          501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1545 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1604

Qy          521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1605 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1664

Qy          541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1665 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1724

Qy          561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1725 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1784

Qy          581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1785 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1844

Qy          601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1845 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1904

Qy          621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1905 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1964

Qy          641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1965 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2024

Qy          661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2025 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2084

Qy          681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2085 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2144

Qy          701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
                |||||      |||||||||||||||||||||||||||||||||||||||||||||||||
Db         2145 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2204

Qy          721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2205 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2264

Qy          741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2265 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2324

Qy          761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2325 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2384

Qy          781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2385 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2444

Qy          801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2445 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2504

Qy          821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2505 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2564

Qy          841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
                |||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db         2565 GCTTCTGATGCTATCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2624

Qy          861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2625 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2684

Qy          881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2685 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2744

Qy          901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2745 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2804

Qy          921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2805 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2864

Qy          941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2865 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2924

Qy          961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2925 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2984

Qy          981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2985 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3044

Qy         1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3045 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3104

Qy         1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3105 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3164

Qy         1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3165 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3224

Qy         1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3225 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3284

Qy         1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3285 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3344

Qy         1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3345 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3404

Qy         1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3405 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3464

Qy         1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3465 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3524

Qy         1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3525 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3584

Qy         1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3585 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3644

Qy         1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro--------- 1217
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3645 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3704
```

```
Qy       1218 ----------------------------------------LeuAlaValValGlnPro 1223
                                                      ||||||||||||||||||
Db       3705 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3764

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3765 ACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3824

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3825 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3884

Qy       1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3885 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3944

Qy       1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3945 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4004

Qy       1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4005 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4064

Qy       1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4065 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4124

Qy       1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4125 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4184

Qy       1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4185 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4244

Qy       1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4245 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4304

Qy       1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4305 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4364

Qy       1424 MetGluLeuTrpLeu 1428
              |||||||||||||||
Db       4365 ATGGAACTTTGGCTG 4379
```

RESULT 12

```
AAD49436
ID   AAD49436 standard; DNA; 4146 BP.
XX
AC   AAD49436;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human blood cell surface antigen, CD109 encoding DNA #1.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; blood cell syrface antigen; CD109; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS.            1..4146
FT                   /*tag=  a
FT                   /product= "Human CD109 protein"
FT                   /transl_except= (pos:2044..2046, aa:Xaa)
FT                   /note= "Xaa corresponds to Ser, Tyr; No start
FT                   and stop codon"
FT                   /partial
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   P-PSDB; AAE32014.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 109-112; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human blood cell
CC   surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.
XX
SQ   Sequence 4146 BP; 1260 A; 776 C; 885 G; 1224 T; 1 other;

Alignment Scores:
Pred. No.:              0           Length:        4146
Score:                  7081.00     Matches:       1379
Percent Similarity:     99.86%      Conservative:  1
Best Local Similarity:  99.78%      Mismatches:    2
Query Match:            96.37%      Indels:        0
DB:                     25          Gaps:          0

US-10-020-095-4 (1-1428) x AAD49436 (1-4146)

Qy         22 ValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGlyAsn 41
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGAAAT 60

Qy         42 ValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAlaGlu 61
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCGGAG 120

Qy         62 LeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPheGlu 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTTGAA 180

Qy         82 LysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIle 101
              |||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||
Db        181 AAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAGATT 240

Qy        102 TyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArg 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACCCGC 300

Qy        122 LeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLys 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATACAAG 360

Qy        142 ProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyrLys 161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTACAAA 420
```

```
Qy        162 ThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSer 181
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTGTCA 480

Qy        182 GlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeu 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATACTT 540

Qy        202 GlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnVal 221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAGGTT 600

Qy        222 SerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMet 241
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCTATG 660

Qy        242 AsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLys 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTGAAA 720

Qy        262 GlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIleThr 281
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATTACA 780

Qy        282 LysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLysAsn 301
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAAAT 840

Qy        302 ValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyProVal 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCAGTA 900

Qy        322 GluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThrAsn 341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACTAAT 960

Qy        342 ValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeuLys 361
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTGAAG 1020

Qy        362 ProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeuThr 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTGACT 1080

Qy        382 LeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGluTyr 401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAGTAC 1140

Qy        402 TrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyrThr 421
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTATACT 1200

Qy        422 ValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGluLeu 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAGCTA 1260

Qy        442 GlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPheLys 461
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTTAAG 1320

Qy        462 SerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGlySer 481
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGATCG 1380

Qy        482 ProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetValVal 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTAGTA 1440

Qy        502 SerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThrPro 521
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACACCA 1500

Qy        522 GluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGlyGlu 541
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGGGAA 1560

Qy        542 IleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLysLeu 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAGCTA 1620

Qy        562 TyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerValThr 581
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTGACA 1680

Qy        582 GlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAla 601
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAATGCC 1740

Qy        602 SerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyr 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGATAT 1800

Qy        622 TyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeu 641
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTATTG 1860

Qy        642 ThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyrAla 661
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATATGCT 1920

Qy        662 GluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeuGly 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTGGGT 1980

Qy        682 SerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMet 701
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATG 2040

Qy        702 GlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrpVal 721
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTG 2100

Qy        722 AlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGlu 741
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAG 2160

Qy        742 LeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGly 761
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGT 2220

Qy        762 GluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLys 781
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAG 2280

Qy        782 ValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAla 801
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db       2281 GTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAATGCC 2340

Qy        802 ThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhePro 821
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTTCCC 2400

Qy        822 IleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAla 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 ATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACTGCT 2460

Qy        842 SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyrSer 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATATTCA 2520

Qy        862 GlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSer 881
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGT 2580

Qy        882 PheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGly 901
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2581 TTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGA 2640

Qy          902 AspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCys 921
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2641 GATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGT 2700

Qy          922 GlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLys 941
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2701 GGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAA 2760

Qy          942 LysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTyr 961
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2761 AAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTAC 2820

Qy          962 GlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAsp 981
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2821 CAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGAC 2880

Qy          982 ProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspPro 1001
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2881 CCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCT 2940

Qy         1002 TyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGln 1021
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2941 TACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAG 3000

Qy         1022 LysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGly 1041
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3001 AAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGC 3060

Qy         1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3061 AATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAG 3120

Qy         1062 TyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSerArg 1081
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3121 TATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGA 3180

Qy         1082 GlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer 1101
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3181 GGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGT 3240

Qy         1102 ProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMet 1121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3241 CCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATG 3300

Qy         1122 GlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAsp 1141
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3301 CAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGAT 3360

Qy         1142 IleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGly 1161
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3361 ATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGA 3420

Qy         1162 IleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThr 1181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3421 ATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACT 3480

Qy         1182 GlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGlu 1201
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3481 CAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAA 3540

Qy         1202 ArgThrAsnIleGlnValThrValThrGlyProSerSerProSerProLeuAlaValVal 1221
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3541 AGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTGGTA 3600

Qy         1222 GlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsn 1241
                ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3601 CAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAAT 3660

Qy         1242 ValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGln 1261
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3661 GTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAA 3720

Qy         1262 GluAlaPheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAsp 1281
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3721 GAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGAT 3780

Qy         1282 LeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluVal 1301
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3781 TTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTT 3840

Qy         1302 AsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLys 1321
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3841 AACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAG 3900

Qy         1322 LysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGln 1341
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3901 AAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAG 3960

Qy         1342 PheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSer 1361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3961 TTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCA 4020

Qy         1362 ValSerIleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGlu 1381
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4021 GTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAA 4080

Qy         1382 ValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAsp 1401
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4081 GTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGAT 4140

Qy         1402 GlyAla 1403
                ||||||
Db         4141 GGAGCT 4146

RESULT 13
AAD49437
ID   AAD49437 standard; DNA; 4197 BP.
XX
AC   AAD49437;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 DNA #2.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4197
FT                   /*tag=  a
FT                   /product= "Human r1520 protein #3"
FT                   /transl_except= (pos:2044..2046, aa:Xaa)
FT                   /note= "Xaa corresponds to Ser, Tyr; No start
FT                   and stop codon"
FT                   /partial
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
```

```
DR   P-PSDB; AAE32015.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 118-120; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human r150 protein
CC   encoding DNA.
XX
SQ   Sequence 4197 BP; 1276 A; 788 C; 894 G; 1238 T; 1 other;

Alignment Scores:
Pred. No.:              0          Length:        4197
Score:                  7068.50    Matches:       1380
Percent Similarity:     98.71%     Conservative:  1
Best Local Similarity:  98.64%     Mismatches:    1
Query Match:            96.20%     Indels:        17
DB:                     25         Gaps:          1

US-10-020-095-4 (1-1428) x AAD49437 (1-4197)

Qy        22 ValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGlyAsn 41
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1 GTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGAAAT 60

Qy        42 ValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAlaGlu 61
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        61 GTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCGGAG 120

Qy        62 LeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPheGlu 81
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 CTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTTGAA 180

Qy        82 LysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIle 101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 AAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAGATT 240

Qy       102 TyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArg 121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 TATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACCCGC 300

Qy       122 LeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLys 141
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 TTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATACAAG 360

Qy       142 ProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyrLys 161
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 CCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTACAAA 420

Qy       162 ThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSer 181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 ACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTGTCA 480

Qy       182 GlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeu 201
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 CAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATACTT 540

Qy       202 GlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnVal 221
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 GGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAGGTT 600

Qy       222 SerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMet 241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 TCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCTATG 660

Qy       242 AsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLys 261
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 AATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTGAAA 720

Qy       262 GlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIleThr 281
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 GGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATTACA 780

Qy       282 LysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLysAsn 301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       781 AAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAAAT 840

Qy       302 ValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyProVal 321
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 GTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCAGTA 900

Qy       322 GluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThrAsn 341
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 GAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACTAAT 960

Qy       342 ValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeuLys 361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 GTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTGAAG 1020

Qy       362 ProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeuThr 381
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 CCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTGACT 1080

Qy       382 LeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGluTyr 401
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 CTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAGTAC 1140

Qy       402 TrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyrThr 421
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 TGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTATACT 1200

Qy       422 ValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGluLeu 441
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1201 GTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAGCTA 1260

Qy       442 GlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPheLys 461
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 CAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTTAAG 1320

Qy       462 SerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGlySer 481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 TCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGATCG 1380

Qy       482 ProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetValVal 501
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 CCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTAGTA 1440

Qy       502 SerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThrPro 521
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1441 TCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACACCA 1500

Qy       522 GluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGlyGlu 541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 GAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGGGAA 1560

Qy       542 IleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLysLeu 561
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 ATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAGCTA 1620

Qy       562 TyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerValThr 581
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 TATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTGACA 1680

Qy       582 GlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAla 601
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 CAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAATGCC 1740
```

```
Qy    602 SerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyr 621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGATAT 1800

Qy    622 TyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeu 641
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 TATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTATTG 1860

Qy    642 ThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyrAla 661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 ACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATATGCT 1920

Qy    662 GluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeuGly 681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTGGGT 1980

Qy    682 SerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMet 701
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 AGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATG 2040

Qy    702 GlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrpVal 721
          |||      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTG 2100

Qy    722 AlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGlu 741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAG 2160

Qy    742 LeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGly 761
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 CTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGT 2220

Qy    762 GluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLys 781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAG 2280

Qy    782 ValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAla 801
          |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db   2281 GTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAATGCC 2340

Qy    802 ThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhePro 821
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 ACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTTCCC 2400

Qy    822 IleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAla 841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 ATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACTGCT 2460

Qy    842 SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyrSer 861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 TCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATATTCA 2520

Qy    862 GlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSer 881
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 CAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGT 2580

Qy    882 PheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGly 901
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 TTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGA 2640

Qy    902 AspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCys 921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGT 2700

Qy    922 GlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLys 941
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAA 2760

Qy    942 LysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTyr 961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 AAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTAC 2820

Qy    962 GlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAsp 981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 CAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGAC 2880

Qy    982 ProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspPro 1001
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCT 2940

Qy   1002 TyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGln 1021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 TACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAG 3000

Qy   1022 LysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGly 1041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 AAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGC 3060

Qy   1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 AATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAG 3120

Qy   1062 TyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSerArg 1081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 TATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGA 3180

Qy   1082 GlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer 1101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 GGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGT 3240

Qy   1102 ProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMet 1121
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 CCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATG 3300

Qy   1122 GlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAsp 1141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGAT 3360

Qy   1142 IleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGly 1161
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 ATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGA 3420

Qy   1162 IleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThr 1181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 ATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACT 3480

Qy   1182 GlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGlu 1201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 CAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAA 3540

Qy   1202 ArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro------------ 1217
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 AGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTTCTG 3600

Qy   1218 ---------------------------------------LeuAlaValValGlnProMet 1224
                                                 |||||||||||||||||||||
Db   3601 ATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCAATG 3660

Qy   1225 AlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyr 1244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 GCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATAT 3720

Qy   1245 AsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAlaPhe 1264
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 AATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTT 3780

Qy   1265 AspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsnVal 1284
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 GATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTG 3840

Qy   1285 CysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeu 1304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTA 3900

Qy   1305 SerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysValGlu 1324
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3901 AGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAA 3960

Qy         1325 TyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysVal 1344
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3961 TATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTT 4020

Qy         1345 AsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIle 1364
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4021 AATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATA 4080

Qy         1365 ValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeu 1384
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4081 GTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTG 4140

Qy         1385 SerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4141 TCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4197

RESULT 14
ABQ79968
ID   ABQ79968 standard; cDNA; 3535 BP.
XX
AC   ABQ79968;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K15 protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K15; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K15"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82169.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 4a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K15 cDNA sequence.
XX
SQ   Sequence 3535 BP; 1108 A; 675 C; 731 G; 1021 T; 0 other;

Alignment Scores:
Pred. No.:              0             Length:        3535
Score:                  5470.00       Matches:       1062
Percent Similarity:     100.00%       Conservative:  1
Best Local Similarity:  99.91%        Mismatches:    0
Query Match:            74.44%        Indels:        0
DB:                     24            Gaps:          0

US-10-020-095-4 (1-1428) x ABQ79968 (1-3535)

Qy            1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy           21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy           41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy           61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy           81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy          101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy          121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy          141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy          161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy          181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy          201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy          221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy          241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy         261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy         281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy         301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy         321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy         341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy         361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy         381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy         401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy         421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy         441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy         461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy         481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy         501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy         521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy         541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy         561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy         581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy         601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy         621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy         641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy         661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy         681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy         701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy         721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy         741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy         761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy         781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
                 ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy         801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy         821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy         841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy         861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy         881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy         901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy         921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy         941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy         961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052
```

```
Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       1061 LysTyrGln 1063
              |||||||||
Db       3293 AAGTATCAG 3301

RESULT 15
ABQ79969
ID   ABQ79969 standard; cDNA; 3535 BP.
XX
AC   ABQ79969;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K15 variant protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K15; variant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K15 variant"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82170.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 4b; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K15 variant cDNA sequence.
XX
SQ   Sequence 3535 BP; 1107 A; 676 C; 731 G; 1021 T; 0 other;

Alignment Scores:
Pred. No.:              0              Length:        3535
Score:                  5461.00        Matches:       1061
Percent Similarity:     99.91%         Conservative:  1
Best Local Similarity:  99.81%         Mismatches:    1
Query Match:            74.32%         Indels:        0
DB:                     24             Gaps:          0

US-10-020-095-4 (1-1428) x ABQ79969 (1-3535)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||.
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832
```

```
Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAAGCCAGTG 892

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy        681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy        701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
Db       2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy        721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy        741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy        761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy        781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy        801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy        821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy        841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy        861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy        881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy        901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy        921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       1061 LysTyrGln 1063
              |||||||||
Db       3293 AAGTATCAG 3301
```

Search completed: January 17, 2004, 01:15:04
Job time : 983 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         January 16, 2004, 21:02:03 ; Search time 198 Seconds
                                             (without alignments)
                                             3183.309 Million cell updates/sec

Title:          US-10-020-095-4
Perfect score:  7348
Sequence:       1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters:      1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10020095/runat_15012004_163823_18719/app_query.fasta_1.1607
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10020095_@CGN_1_1_118_@runat_15012004_163823_18719 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1441	19.6	4577	4	US-09-241-606-1	Sequence 1, Appli
	2	1346.5	18.3	4079	4	US-09-016-434-1174	Sequence 1174, Ap
	3	973.5	13.2	5211	1	US-08-447-411-1	Sequence 1, Appli
	4	931	12.7	5056	2	US-08-793-126-2	Sequence 2, Appli
	5	931	12.7	5056	3	US-09-132-271-2	Sequence 2, Appli
	6	931	12.7	5067	3	US-09-142-334-23	Sequence 23, Appl
	7	894	12.2	5948	2	US-08-662-227-1	Sequence 1, Appli
	8	894	12.2	5948	4	US-09-017-947-1	Sequence 1, Appli
	9	885	12.0	5924	1	US-08-447-411-44	Sequence 44, Appl
	10	768.5	10.5	4138	1	US-08-447-411-75	Sequence 75, Appl
	11	768.5	10.5	4138	2	US-08-662-227-33	Sequence 33, Appl
	12	768.5	10.5	4138	4	US-09-017-947-33	Sequence 33, Appl
c	13	245	3.3	18431	4	US-09-221-017B-1090	Sequence 1090, Ap
	14	234	3.2	339	4	US-09-311-352B-1	Sequence 1, Appli
	15	215.5	2.9	750	4	US-09-241-606-3	Sequence 3, Appli
c	16	203.5	2.8	14066	4	US-09-601-198-56	Sequence 56, Appl
	17	185	2.5	549	4	US-09-311-352B-3	Sequence 3, Appli
	18	180	2.4	4026	4	US-09-252-991A-10214	Sequence 10214, A
	19	174.5	2.4	7101	1	US-08-480-604A-9	Sequence 9, Appli
	20	174.5	2.4	7101	2	US-08-405-496A-9	Sequence 9, Appli
	21	174.5	2.4	7101	3	US-08-915-136-9	Sequence 9, Appli
	22	174.5	2.4	7101	4	US-08-957-310-9	Sequence 9, Appli
	23	174.5	2.4	7101	4	US-10-011-366-9	Sequence 9, Appli
	24	171	2.3	9873	4	US-09-328-352-1360	Sequence 1360, Ap
	25	165	2.2	28473	4	US-08-961-527-83	Sequence 83, Appl
	26	164.5	2.2	4977	4	US-09-071-035-257	Sequence 257, App
	27	164.5	2.2	4977	4	US-09-071-035-261	Sequence 261, App
	28	164.5	2.2	4977	4	US-09-071-035-265	Sequence 265, App
	29	164.5	2.2	13993	4	US-09-220-132-20	Sequence 20, Appl
	30	163.5	2.2	14070	4	US-09-108-006C-2	Sequence 2, Appli
	31	163	2.2	37948	3	US-09-251-645-11	Sequence 11, Appl
c	32	161	2.2	3312	4	US-09-601-198-58	Sequence 58, Appl
	33	161	2.2	7286	4	US-08-793-273C-1	Sequence 1, Appli
	34	161	2.2	7286	5	PCT-US95-11684-1	Sequence 1, Appli
c	35	160	2.2	5629	4	US-09-453-702B-243	Sequence 243, App
c	36	157.5	2.1	15016	4	US-09-601-198-60	Sequence 60, Appl
c	37	156	2.1	3636	4	US-08-961-527-78	Sequence 78, Appl
	38	156	2.1	30549	4	US-09-134-001C-322	Sequence 322, App
	39	155.5	2.1	9510	4	US-09-453-702B-256	Sequence 256, App
	40	154	2.1	9171	1	US-08-038-682-5	Sequence 5, Appli
	41	154	2.1	9171	1	US-08-302-832-5	Sequence 5, Appli
	42	154	2.1	9171	2	US-08-530-198-5	Sequence 5, Appli
	43	154	2.1	9171	2	US-08-469-880-5	Sequence 5, Appli
	44	154	2.1	9171	2	US-08-728-470-5	Sequence 5, Appli
	45	154	2.1	9171	2	US-08-617-697-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-09-241-606-1
; Sequence 1, Application US/09241606
; Patent No. 6472140
; GENERAL INFORMATION:
;  APPLICANT: Tanzi, Rudolph E.
;  APPLICANT: Kovacs, Dora
;  APPLICANT: Saunders, Aleister J.
;  TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
;  TITLE OF INVENTION: Alzheimer's Disease
;  FILE REFERENCE: 0609.4460003
;  CURRENT APPLICATION NUMBER: US/09/241,606
;  CURRENT FILING DATE: 1999-02-02
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 4577
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: (44)..(112)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (44)..(4468)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (113)..(4468)
US-09-241-606-1

Alignment Scores:
Pred. No.:              1.49e-142      Length:         4577
Score:                  1441.00        Matches:        451
Percent Similarity:     46.20%         Conservative:   266
Best Local Similarity:  29.06%         Mismatches:     551
Query Match:            19.61%         Indels:         284
```

```
DB:                    4        Gaps:        59
US-10-020-095-4 (1-1428) x US-09-241-606-1 (1-4577)

Qy         12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db         86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy         32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
Db        146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy         52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
Db        203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy         72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
Db        263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy         90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
Db        311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db        365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
Db        416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy        147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db        476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy        167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
Db        536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy        186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db        596 GGTGGCCTCAAGCAATTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy        206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db        656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy        224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db        713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772

Qy        244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db        773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGGAAGCCTGTCCCTGGACAT 832

Qy        264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
Db        833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy        272 ---------------------------------------------SerPheTrpGlyLys 276
Db        893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952

Qy        277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db        953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 988

Qy        297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
Db        989 TATGAAATGAAA------------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy        317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
Db       1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy        332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db       1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy        352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
Db       1133 CCCTTCTTTGGG---------------------------------------------CAG 1147

Qy        372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
Db       1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1195

Qy        392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
Db       1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy        412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
Db       1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy        429 LysIleGluPhePro------------------------------IleLeuGluAspSer 438
Db       1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy        439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db       1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390

Qy        459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db       1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy        479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492
Db       1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy        493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
Db       1511 CTGAAGAAGCTCTCCTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy        511 ------------------LysGlnAsnSerThr------------MetPheSerLeuThr 520
Db       1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db       1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1690

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db       1691 GACGTGATTGGGGATTCTGCAAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db       1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db       1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
Db       1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927

Qy        620 GlyTyr------------------------------------------TyrLeuGlyMet 625
Db       1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987

Qy        626 PheMetAsnSer------------------------------------------PheAla 631
```

```
Db   1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047

Qy    632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
Db   2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107

Qy    652 AspGlyValTyrAspAsnAlaGlu--------------------------------Tyr 660
Db   2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167

Qy    661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db   2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218

Qy    681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
Db   2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269

Qy    698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
Db   2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329

Qy    718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
Db   2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389

Qy    738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
Db   2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446

Qy    758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
Db   2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506

Qy    778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
Db   2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566

Qy    798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db   2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA--------------------- 2605

Qy    813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db   2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy    833 ThrValThrAlaLeuSer--------------------------------ProThrAla 841
Db   2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy    842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
Db   2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy    857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
Db   2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy    877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
Db   2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy    897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
Db   2885 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy    917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
Db   2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy    937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
Db   3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGATCAAGTCCAAGGCCATTGGCTAT 3064

Qy    957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
Db   3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy    977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
Db   3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy    995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
Db   3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244

Qy   1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
Db   3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy   1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db   3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy   1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
Db   3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy   1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db   3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy   1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
Db   3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy   1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
Db   3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy   1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
Db   3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy   1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
Db   3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy   1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
Db   3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy   1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
Db   3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy   1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
Db   3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy   1216 SerProLeuAlaValValGln--------------------------------------- 1222
Db   3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy   1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
Db   3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy   1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
Db   4016 CAGACCTCCTTGAAATACAATATT---------------------------------CTC 4042

Qy   1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
Db   4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102
```

```
Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
Db       4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
Db       4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
Db       4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
Db       4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
Db       4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
Db       4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438
```

```
RESULT 2
US-09-016-434-1174
; Sequence 1174, Application US/09016434
; Patent No. 6500938
;  GENERAL INFORMATION:
;    APPLICANT:  Janice Au-Young
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF SIGNALING
;    TITLE OF INVENTION:  PATHWAY GENE EXPRESSION
;    NUMBER OF SEQUENCES:  1490
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,434
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0002 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  1174:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4079 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GENBANK
;      CLONE:  g177869
US-09-016-434-1174

Alignment Scores:
Pred. No.:              1.46e-132       Length:         4079
Score:                  1346.50         Matches:        419
Percent Similarity:     46.66%          Conservative:   238
Best Local Similarity:  29.76%          Mismatches:     488
Query Match:            18.32%          Indels:         263
DB:                     4               Gaps:           54

US-10-020-095-4 (1-1428) x US-09-016-434-1174 (1-4079)
```

```
Qy         12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db         86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy         32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
Db        146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy         52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
Db        203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy         72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
Db        263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy         90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
Db        311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db        365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
Db        416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy        147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db        476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy        167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
Db        536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy        186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db        596 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy        206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db        656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy        224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db        713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772

Qy        244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db        773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGGAAGCCTGTCCCTGGACAT 832

Qy        264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
Db        833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy        272 ---------------------------------------------SerPheTrpGlyLys 276
Db        893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952

Qy        277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db        953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 988
```

```
Qy        297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
              |||||||||                        |||            :::
Db        989 TATGAAATGAAA----------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy        317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
                 |||   |||    ||||||                     :::|||   ::::::
Db       1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy        332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
              |||   :::|||     |||      :::      |||:::|||            |||
Db       1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy        352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
                 ||||||                                                :::
Db       1133 CCCTTCTTTGGG---------------------------------------------CAG 1147

Qy        372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
              |||         ||||||      :::   :::                |||   ||||||
Db       1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1195

Qy        392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
              :::   :::   |||       |||:::      ||||||:::               :::
Db       1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy        412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
                       ||||||      ||||||   |||      :::      |||   |||
Db       1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy        429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438
              ::::::   :::                                :::   |||:::
Db       1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy        439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
                 |||            ||||||
Db       1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390

Qy        459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
              |||      ||||||||||||:::::::::   |||:::               :::
Db       1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy        479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492
                 |||                         :::::::::|||
Db       1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy        493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
              ||||||:::||||||      |||:::::::::::::::|||   :::|||      |||
Db       1511 CTGAAGAAGCTCTCCTTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy        511 ------------------LysGlnAsnSerThr-----------MetPheSerLeuThr 520
                                |||||                     |||:::
Db       1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ::::::         |||   |||   :::::::::   :::                |||
Db       1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1690

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ::::::|||   |||      |||      |||:::            ||||||:::
Db       1691 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              |||   :::|||      ::::::   |||:::            ||||||::::::
Db       1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
                 |||   |||:::      :::   |||||||||:::||||||   ||||||
Db       1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
                 :::::::::      :::|||   :::   |||   |||            |||
Db       1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927

Qy        620 GlyTyr-----------------------------------------TyrLeuGlyMet 625
              |||:::                                               :::
Db       1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987

Qy        626 PheMetAsnSer-----------------------------------------PheAla 631
              ::::::|||                                            ::::::
Db       1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047

Qy        632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
                 ::::::      ||||||      |||::::::   :::   |||      :::
Db       2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107

Qy        652 AspGlyValTyrAspAsnAlaGlu--------------------------------Tyr 660
                                 |||                                 |||
Db       2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
                    |||                     :::|||   :::   :::
Db       2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218

Qy        681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
                    ||||||         |||||||||:::||||||||||||||||||||
Db       2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269

Qy        698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
                    :::               |||      |||||||||||||||:::|||
Db       2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329

Qy        718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
              |||   |||   |||      |||   :::||||||||   |||||||||:::::::::
Db       2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389

Qy        738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
              |||      |||:::|||||||||||||||||||:::   |||   :::|||||||||
Db       2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446

Qy        758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
              |||||||||||||||   |||   |||:::   |||:::   |||||||||
Db       2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506

Qy        778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
                 |||   |||   :::|||   |||         |||         :::
Db       2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566

Qy        798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
              |||               |||   |||   |||   |||
Db       2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA-------------------- 2605

Qy        813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
                       ||||||   :::   :::   |||      ||||||   :::
Db       2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy        833 ThrValThrAlaLeuSer--------------------------------ProThrAla 841
              ||||||:::|||   :::                                |||:::
Db       2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGACTGAGGTGCCTTCAGTT 2713

Qy        842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
                             |||   |||   :::   :::||||||:::   ||||||:::
Db       2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy        857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
              ||||||      :::   :::   ||||||               :::
Db       2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy        877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
                 :::   ||||||         |||||||||   |||      |||   |||
Db       2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy        897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
              ::::::   :::|||||||:::||||||      ::::::         :::|||::::::
Db       2885 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy        917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
              |||||||||||||||||||||||||||:::      ||||||||||||||||:::|||
```

```
Db      2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy       937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
             ||||||||    :::    :::||||||||      :::|||   |||||:::    :::
Db      3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3064

Qy       957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
             :::      ||||||||||||:::|||   |||:::      |||||||||:::|||
Db      3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy       977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
             ||||||       |||     :::      |||:::||||||||:::||||||||||||:::
Db      3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy       995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
                |||   :::|||      ||||||   ||||||:::         :::   :::
Db      3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244

Qy      1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
                ||||||          ||||||   ||||||   |||          |||   ::::::
Db      3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy      1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
             ::::::   ::::::||||||   :::         |||||||||:::|||||||||
Db      3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy      1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
             :::||||||                     :::      |||:::   ::::::
Db      3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy      1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
             |||||||||                  |||      |||   |||      ||||||
Db      3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy      1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
                ||||||:::   ||||||   :::      |||:::      |||   ||||||
Db      3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy      1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
             |||      |||         |||   ::::::         |||
Db      3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy      1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
                   ::::::   :::         ::::::|||:::         |||   ::::::
Db      3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy      1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
             |||:::   :::|||   ||||||:::   :::|||      |||         ||||||
Db      3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy      1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
             |||   :::               |||::::::|||:::::::::|||:::|||:::
Db      3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy      1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
             |||||||||:::||||||||||||||||   |||||||||   |||||||||::::::
Db      3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy      1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
             |||         |||   ||||||            ||||||      ||||||
Db      3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy      1216 SerProLeuAlaValValGln--------------------------------------- 1222
             :::      :::   |||
Db      3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy      1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
                         |||         ::::::         |||   |||
Db      3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy      1239 GlnLeuAsnValValTyrAsnVal 1246
             |||   ::::::   ||||||:::
Db      4016 CAGACCTCCTTGAAATACAATATT 4039
```

```
RESULT 3
US-08-447-411-1
; Sequence 1, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5211 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  9..4961
US-08-447-411-1

Alignment Scores:
Pred. No.:              1.37e-92      Length:         5211
Score:                  973.50        Matches:        392
Percent Similarity:     40.56%        Conservative:   280
Best Local Similarity:  23.66%        Mismatches:     650
Query Match:            13.25%        Indels:         335
DB:                     1             Gaps:           62

US-10-020-095-4 (1-1428) x US-08-447-411-1 (1-5211)

Qy         1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
             |||::::|||         |||         |||   ||||||            :::
Db         9 ATGGAGGGGATGGCTCTCTATCTGGTGGCTGCTCTATTGATTGGTTTTCCAGGGTCT--- 65

Qy        21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
                   :::   |||      :::   :::      |||   ::::::|||
Db        66 ------TCCCACGGGGCTCTCTATACCCTCATCACCCCTGCTGTTTTGCGAACAGACACA 119

Qy        41 AsnValThrIleGlyValGluLeuLeu---GluHisCysProSerGlnValThrValLys 59
                   |||   ||||||         :::      |||      :::   :::
Db       120 GAAGAGCAAATTTTGGTGGAGGCCCATGGAGACAGTACTCCAAAATCGCTTGACATCTTT 179
```

```
Qy     60 Ala------------GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGlu 75
Db    180 GTTCATGATTTTCCACGGAAGCAGAAAACCTTGTTCCAAAGCAGAGTAGATATGAATCAG 239

Qy     76 AlaGluGlyValPheGluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeu 95
Db    240 GCAGGAAGCATGTTTGTCACT---------CCAACTATAAAGGTTCCTGCAAAAGAACTG 290

Qy     96 AsnSerAlaAspGluIleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeu 115
Db    291 AATAAGGACTCCAAGCAAAATCAGTATGTGGTTGTGAAAGTAACTGGTCCTCAAGTGGCA 350

Qy    116 PheSerAsnSerThrArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThr 135
Db    351 TTGGAAAAGGTGGTTCTCCTTTCTTACCAGAGTGGC------TTTGTGTTCATCCAGACA 404

Qy    136 AspLysAlaLeuTyrLysProLysGlnGluValLysPheArgIleValThrLeuPheSer 155
Db    405 GATAAAGGCATCTATACACCAGGCTCTCCAGTGCGTTATCGTGTCTTTTCTGTG------ 458

Qy    156 AspPheLysProTyrLysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeu 175
Db    459 GATCACAACATGCACAGGATGGACAAAACTGTGATTGTCGAG------------------ 500

Qy    176 IleGlnGlnTrpLeuSerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeu 195
Db    501 ------------------TTTCAGACTCCAGAAGGCATTGTT---------------GTC 527

Qy    196 SerSerHisProIle--------------------------------------------- 200
Db    528 AGTTCTAAACCAGTCAATCCATCAGGCTCGATCCGGCCTTACAATTTACCAGAGCTTGTC 587

Qy    201 ---LeuGlyAspTrpSerIleGlnValGlnValAsn---------AspGlnThrTyrTyr 216
Db    588 AGTTTTGGGACATGGAAGGCTGTGGCCAAATATGAACATTCACCAGAAGAGAGCTACACT 647

Qy    217 GlnSerPheGlnValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrPro 236
Db    648 GCATATTTTGATGTCAGAGAATATGTGTTACCAAGCTTTGAAGTCCGTCTGCAACCATCA 707

Qy    237 ---------LeuTyrCysSerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLys 253
Db    708 GATAAGTTTCTTTACATTGATGGGAAT---AAAAATTTCCACGTGTCTATCACTGCAAGG 764

Qy    254 TyrThrTyrGlyLysProValLysGlyAspValThrLeuThrPheLeuProLeuSerPhe 273
Db    765 TACTTATATGGAAAGAAAGTGGAAGGT---GTGGCCTTTGTCGTCTTTGGAGTCAAAATA 821

Qy    274 TrpGlyLysLysLysAsnIle------------ThrLysThrPheLysIleAsnGlySer 289
Db    822 GATGATGCTAAAAAGAGTATTCCAGACTCACTCACGAGAATTCCGATTATTGATGGAGAT 881

Qy    290 AlaAsnPheSerPheAsnAspGluGluMetLysAsnValMetAspSerSerAsnGlyLeu 309
Db    882 GGGGAAGCAACACTAAAAAGAGATACACTACGTTCCCGATTTCAAGATCTCAATCAGCTT 941

Qy    310 SerGluTyrLeuAspLeuSerSerProGlyProValGluIleLeuThrThrValThrGlu 329
Db    942 GTTGGTCATACTCTGTATGTATCT------------------GTAACAGTGATAACAGAA 983

Qy    330 SerValThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyr 349
Db    984 TCAGGCAGTGATATGGTAGTGACTGAGCAAGGCGGCATTCATATTGTGACATCTCCCTAT 1043

Qy    350 IleIleGluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThr 369
Db   1044 CAGATCTACTTCACAAAAACCCCCAAATATTTCAAGCCAGGAATGCCATATGAACTGACG 1103

Qy    370 ValLysValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnVal 389
Db   1104 GTGTATGTTACCAACCCTGATGGCTCACCAGCTGCC---------------------CAT 1142

Qy    390 ValIleThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsn 409
Db   1143 GTGCCAGTGGTATCAGAGGCCATTCATTCTGAG---------------------GGAACC 1181

Qy    410 GlnLysMetGluAlaValGlnLysIleAsnTyrThrValPro---------GlnSerGly 426
Db   1182 ACTTTGAGTGATGGGACTGCTAAGCTCATTCTGAACACACCACTGAACATTCAAAGC--- 1238

Qy    427 ThrPheLysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyr 446
Db   1239 ---------------CTACCGATCACTGTTAGAACTAACCATGGAGACCTCCCAAGAGAA 1283

Qy    447 PheLeuGlySerLysSerSerMetAlaValHisSerLeuPheLysSerProSerLysThr 466
Db   1284 CGCCAGGCAATAAAGTCCATGACAGCCACAGCCTACCAAACCCAGGGAGGGTCTGAAAAC 1343

Qy    467 TyrIleGlnLeuLysThrArgAspGluAsnIleLysValGlySer------ProPheGlu 484
Db   1344 TATCTTCATGTAGCCATTACATCTACAGAGATTAAGCCCGGAGATAACTTACCTGTCAAT 1403

Qy    485 LeuValValSerGlyAsn------------LysArgLeuLysGluLeuSerTyrMetVal 500
Db   1404 TTCAATGTGAGGGGCAATGCAAATTCACTGAACCAGATCAAATATTTCACATACCTCATA 1463

Qy    501 ValSerArgGlyGlnLeuValAlaValGlyLys------------------GlnAsnSer 514
Db   1464 TTGAATAAAGGGAAGATTTTCAAGGTTGGCAGGCAACCCAGGAGAGATGGGCAGAATCTG 1523

Qy    515 ThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleValTyr 534
Db   1524 GTGACCATGAATCTGCATATCACTCCAGATCTCATCCCTTCCTTCCGGTTTGTGGCTTAC 1583

Qy    535 TyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGln------ 552
Db   1584 TACCAAGTGGGAAATAACGAAATTGTGGCTGATTCTGTCTGGGTGGATGTGAAGGATACC 1643

Qy    553 ------------LeuValPheLys------------------AsnLysIleLysLeuTyr 562
Db   1644 TGCATGGGAACGTTGGTTGTGAAAGGAGCGTCTTCCAGAGACGATCGAATACAA------ 1697

Qy    563 TrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerValThrGln 582
Db   1698 ------------------AAGCCAGGAGCTGCAATGAAAATCAAATTGGAA---GGGGAT 1736

Qy    583 ProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSer 602
Db   1737 CCAGGTGCTCGGGTTGGTCTTGTGGCTGTGGACAAAGCAGTATATGTTCTCAATGATAAA 1796

Qy    603 AsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyr 622
Db   1797 TATAAGATTAGCCAAGCTAAGATATGGGACACAATAGAAAAGAGTGACTTTGGCTGTACA 1856

Qy    623 LeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeuThr 642
Db   1857 GCTGGCAGTGGCCAGAATAATCTGGGTGTGTTTGAAGATGCTGGACTGGCTCTGACAACC 1916

Qy    643 AspAlaAsnLeuThr--------------------------------------------- 647
Db   1917 AGCACTAATCTCAACACCAAACAGAGATCAGCTGCAAAGTGTCCTCAGCCTGCAAATCGG 1976

Qy    647 ------------------------------------------------------------ 647
Db   1977 AGGCGTCGCAGTTCTGTTTTGCTGCTTGACAGCAAAGCAAGCAAAGCGGCACAGTTTCAG 2036

Qy    648 ---------------LysAspTyrIleAspGlyValTyrAspAsn--------------- 657
Db   2037 GATCAAGGCCTGCGTAAATGCTGTGAAGATGGCATGCATGAGAACCCCATGGGGTACACT 2096

Qy    658 AlaGluTyrAlaGluArgPheMetGluGluAsnGluGly--------------------- 670
Db   2097 TGTGAAAAGCGTGCAAAATACATCCAGGAGGGAGATGCTTGTAAGGCTGCCTTCCTTGAA 2156

Qy    671 ------HisIleVal--------------------------------------------- 673
```

```
Db       2157 TGTTGTCACTACATCAAAGGGATCCGAGATGAAAACCAACGGGAGAGCGAGTTGTTTCTG 2216

Qy        674 ---------AspIleHisAspPheSerLeuGlySerSerProHisVal---ArgLysHis 689
                       |||      |||         |||          :::    |||
Db       2217 GCAAGAAGTGATTTTGAAGATGAACTCTTTGGAGATGACAACATCATCTCCAGGTCTGAT 2276

Qy        690 PheProGluThrTrpIleTrpLeu--------------------AspThrAsnMetGly 702
              ||||||||::||:::||||||                        |||   |||
Db       2277 TTTCCCGAGAGTTGGTTGTGGCTAACAGAGGAATTGACCGGGGAGCCTAACAATCAAGGG 2336

Qy        703 TyrArgIleTyrGlnGluPheGluValThrValPro------------AspSerIleThr 718
                                       |||||||||               ||||||||||||
Db       2337 ATTTCA---------------AGCAAGACAGTACCTTTTATCTGAGGGATTCCATCACA 2381

Qy        719 SerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThr 738
              :::|||                         |||   |||   :::||||||:::    |||
Db       2382 ACCTGG------------------------GAGTTGCTGGCTGTGGGCCTTTCACCCACC 2417

Qy        739 ---------------------ProValGluLeuGlnAlaPheGlnProPhePheIlePhe 751
                                   |||   |||:::       :::   ||||||||
Db       2418 AAAGGGATCTGTGTGGCTGAACCGTATGAAATAACAGTCATGAAAGACTTCTTCATTGAT 2477

Qy        752 LeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePhe 771
              |||   ||||||||||||||:::::: |||:::      :::        ::::::
Db       2478 CTTCGACTGCCATATTCAGTAGTGAAGAATGAGCAGGTGGAGATTCGAGCTATTCTGTAC 2537

Qy        772 AsnTyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAsp 791
              ||||||      :::        |||:::|||      :::        |||
Db       2538 AACTACGCTGACGAGGATATTTATGTGCGAGTGGAACTGATATACAACCCAGCCTTC--- 2594

Qy        792 IleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValPro 811
                                |||:::          ::::::|||
Db       2595 ---------TGCAGTGCTTCCACAGAAGGACAAAGATACCGACAGCAGTTCCCAATTAAA 2645

Qy        812 SerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly------Glu 829
              :::                |||   |||   |||   |||      |||          :::
Db       2646 GCCCTGTCCTCCAGAGCAGTACCATTTGTGATAGTCCCATTAGAGCAAGGATTGCATGAT 2705

Qy        830 IleProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIle 849
              :::   :::      :::               ||||||||   |||   :::   :::
Db       2706 GTTGAGGTTATAGCAAGTGTCCGGGGAGAGTTGGCATCAGATGGTGTGAGGAAGAAACTG 2765

Qy        850 LeuValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThr 869
              |||      |||||||      |||:::                |||   ||||||   :::
Db       2766 AAAGTTGTACCTGAAGGGGAACGGAAAAAATATTGTGACTATTATTGAACTGGACCCAAGT 2825

Qy        870 AspAsnArgLeuGlnSerThr------LeuLysThrLeuSerPheSerPheProProAsn 887
                    :::         |||   ::::::
Db       2826 GTAAAAGGAGTTGGTGGAACCCAGGAACTAACGGTCATAGCCAATAAATTAGATGACAAG 2885

Qy        888 ThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAsp------------Val 903
                                :::|||:::   :::||||||                  :::
Db       2886 GTGCCTGATACAGAAGTTGAGACCAGGATTTCTGTTCTAGGTGACCCTGTGGCTCAGATT 2945

Qy        904 LeuGlyProSerIleAsnGly------LeuAlaSerLeuIleArgMetProTyrGlyCys 921
              :::      ||||||:::|||      |||      ||||||      |||   ||||||
Db       2946 ATTGAAAACTCAATTGATGGAAGTAAACTCAATCATCTCATTATTACTCCTTCTGGCTGT 3005

Qy        922 GlyGluGlnAsnMetIleAsnPheAlaProAsnIle---------TyrIleLeuAspTyr 938
              ||||||||||||||||||               |||::::::         |||   ||||||
Db       3006 GGGGAGCAAAATATGATCACCATGACTCCATCGGTCATTGCCACCTACTACTTGGACGCA 3065

Qy        939 LeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArg 958
                 :::   :::      |||         ::::||::::          :::
Db       3066 ACAGGGCAGTGGGAGAATCTTGGTGTGGATCGCAGGACTGAAGCTATCAAACAGATCATG 3125

Qy        959 GlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGly 978
                 ||||||   ::::::::::::|||::::::   |||   |||::::::||||||
Db       3126 ACTGGTTATGCCCAGCAGATGGTGTACAAGAAAGCAGATCATTCCTATGCAGCATTTACA 3185

Qy        979 AsnTyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGlu 998
              |||         :::   |||:::||||||:::|||:::|||::::::
Db       3186 AAC------CGTGCATCTAGTTCTTGGCTAACAGCATATGTGGTGAAAGTCTTAGCCATG 3239

Qy        999 AlaAspProTyrIle---AspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeu 1017
              |||      :::      ||||||         ::::::                ||||||
Db       3240 GCTTCCAACATGGTAAAAGACATTAGCCATGAAATTATTTGTGGAGGTGTGAAATGGCTC 3299

Qy       1018 ---LysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSer 1036
                          |||:::   :::|||   |||   :::            ||||||||
Db       3300 ATTCTGAACAGGCAACAACCAGATGGAGTGTTCAAAGAAAATGCCCCTGTGATCCATGGA 3359

Qy       1037 GluLeuGlnGlyGlyAsnLysSer---------ProValThrLeuThrAlaTyrIleVal 1053
              |||:::   ||||||   |||                  :::|||||||||:::||||||
Db       3360 GAAATGCTGGGAGGAACTAAAGGTGCTGAACCAGAAGCATCTTTAACAGCATTCATTGTG 3419

Qy       1054 ThrSerLeuLeuGlyTyrArg---------LysTyrGlnProAsnIle---------Asp 1067
              |||:::||||||        |||            |||   |||   ||||||
Db       3420 ACTGCATTATTGGAATCCAGATCAGTCTGCAAAGAACAAATCAATATTCTAGACAGCAGC 3479

Qy       1068 ValGlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyr 1087
              :::   ::::::      :::|||         ::::::   :::   :::        |||
Db       3480 ATCAATAAGGCCACAGATTATTTACTCAAAAAGTATGAGAAA---CTGCAAAGGCCTTAC 3536

Qy       1088 ThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAla 1107
              |||   ||||||      |||||||||:::                  |||   :::
Db       3537 ACTACAGCCCTCACAGCCTATGCTTTGGCT------------------GCTGCAGACCGA 3578

Qy       1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSer 1127
              ||||||                                 |||         |||
Db       3579 CTCAATGATGACAGGGTACTCATGGCAGCATCAACAGGAAGGAATCGTTGG--------- 3629

Qy       1128 GluSerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyr 1147
                       :::   :::         |||:::   :::||||||         :::|||
Db       3630 ---------------GAAGAATATAATGCTCGCACCCATAATATTGAAGGCACTTCCTAT 3674

Qy       1148 AlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrp 1167
              |||||||||:::          :::   :::|||         |||::::::||||||
Db       3675 GCCTTGTTGGCCCTGCTGAAAATGAAGAAATTTGCTGAGGTCGGTCCTGTAGTCAGATGG 3734

Qy       1168 LeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAla 1187
              |||      |||:::         |||   :::         ||||||   |||   :::
Db       3735 CTGATAGATCAGAAATATTATGGGGAACATATGGACAAACCCAAGCAACAGTTATGGTG 3794

Qy       1188 LeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg---ThrAsnIleGln 1206
                 :::||||||:::|||:::         |||   |||   :::      |||:::
Db       3795 TTTCAAGCTCTTGCTGAATATGAGATTCAGATGCCTACCCATCAGGACTTAAATTTAGAT 3854

Qy       1207 ValThrValThrGlyProSerSerProSerPro--------------------LeuAla 1219
              ::::::::      |||         |||
Db       3855 ATTTCTATTAAACTGCCAGAACGAGAAGTACCTGAAAGGTACAGCATTAATGATAGAAAT 3914

Qy       1220 ValValGlnProMetAlaVal-----------------------AsnIleSerAlaAsn 1231
              ||||||   |||                              :::||||||:::
Db       3915 GCTGTCCAGGCCCGGACAGTAGAGACCAAACTCAACGAAGACTTCACTGTGTCAGCATCA 3974

Qy       1232 GlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySer 1251
              |||   |||   |||         :::      |||||||||   :::
Db       3975 GGTGATGGAAAAGCAACAATGACCATTTTGACGGTCTATAATGCACAATTG--------- 4025

Qy       1252 SerArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaVal--- 1270
                 |||         ::::::   |||:::         |||   |||||||||:::|||
Db       4026 ---AGGGAGGATGCAAATGTTTGCAACAAA------TTCCATCTTGATGTTTCTGTTGAA 4076

Qy       1271 ------------------LysGluAsnLysAspAspLeuAsnHisValAspLeuAsnVal 1284
                                |||:::   |||                :::   |||   :::
Db       4077 AACGTCGAATTGAACTTAAAACAGGCAAAGGGAGGCAAGGCAGCCCTCAGGCTTAAAATC 4136

Qy       1285 CysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeu 1304
              ||||||   :::   |||         |||   |||   ::::::::::::::::::|||
Db       4137 TGCACTAGGTATCTGGGAGAAGTTGATTCTACAATGACAATAATTGATATTTCTATGCTG 4196

Qy       1305 SerGlyPheMetValProSerGluAlaIle---------------SerLeuSerGluThr 1319
              :::||||||            :::|||   :::                   :::
Db       4197 ACTGGTTTTTCCCTGATGCTGAAGACCTTAAAAGGCTTTCTAACGGAGTGGACAGATAC 4256
```

```
Qy       1320 ValLysLysValGluTyrAspHis---------------GlyLysLeuAsnLeuTyrLeu 1334
              :::   |||   |||   |||:::                |||   :::   :::||||||
Db       4257 ATCTCCAAGTTTGAAATTGACAATAATATGGCTCAGAAAGGAACTGTTGTCATTTACTTA 4316

Qy       1335 AspSerValAsnGluThrGln---PheCysValAsnIleProAlaValArgAsnPheLys 1353
              |||   |||:::   ::::::      |||::::::              ::::::|||:::
Db       4317 GACAAGGTCTCCCACTCTGAAGATGAATGCCTGCACTTTAAGATTCACAAGCATTTTGAA 4376

Qy       1354 ValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGln 1373
              |||         |||      ||||||   :::      ||||||            |||
Db       4377 GTTGGCTTCATTCAGCCAGGATCAGTCAAGGTGTACAGCTACTACAATCTAGATGAACAA 4436

Qy       1374 AlaValArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspVal 1393
                    :::   |||:::   :::   :::               :::
Db       4437 TGTACCAAGTTCTACCATCCAGATAAAGAAACAGGTCTTCTCAATAAGATATGTCATGGT 4496

Qy       1394 GlnGlyCysArgProCysGluAspGlyAlaSerGlySerHisHisHis----------- 1409
                    ||||||      |||:::      |||      :::
Db       4497 AACATTTGCCGATGTGCAGAAGAAACCTGTTCCTTGCTCAACCAGCAGAAAAAGATTGAT 4556

Qy       1410 -----------------------------SerSerValIlePhePheIleCysPheLys 1419
                                           :::|||   :::::::::      |||
Db       4557 CTTCAATTACGAATTCAAAAAGCCTGCGCGCAAAATGTGGATTATGTCTACAAAACCAAG 4616

Qy       1420 LeuLeu-----------------------------TyrPheMetGluLeu 1426
              ||||||                             |||||||||::::::
Db       4617 CTGCTTCGAATAGAAGAAAAAGATGGTAATGATATCTATTTCATGGATGTT 4667
```

RESULT 4
US-08-793-126-2
; Sequence 2, Application US/08793126
; Patent No. 5849297
; GENERAL INFORMATION:
; APPLICANT: Harrison, Richard Alexander
; APPLICANT: Farries, Charles Timothy
; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: United States of America
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,126
; FILING DATE: 07-FEB-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 102286.377
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 526-6000
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5056 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-793-126-2

```
Alignment Scores:
Pred. No.:              4.44e-88      Length:          5056
Score:                  931.00        Matches:         381
Percent Similarity:     40.88%        Conservative:    287
Best Local Similarity:  23.32%        Mismatches:      609
Query Match:            12.67%        Indels:          358
DB:                     2             Gaps:            63

US-10-020-095-4 (1-1428) x US-08-793-126-2 (1-5056)

Qy          3 GlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeuAlaVal 22
              ||||||   ||||||               ||||||   :::               ||||||:::
Db         65 GGTCCCAGCCTGCTG-----------CTCCTGCTACTAACCCACCTCCCCCTGGCTCTG 112

Qy         23 AlaProGlyProArgPheLeuValThrAlaProGlyIleIleArg--------------- 37
                    |||   :::   :::      |||   |||:::|||
Db        113 GGG---AGTCCCATGTACTCTATCATCACCCCCAACATCTTGCGGCTGGAGAGCGAGGAG 169

Qy         38 ----------------------ProGlyGlyAsnValThrIleGlyValGluLeuLeu 49
                                    |||:::|||   :::   |||   :::
Db        170 ACCATGGTGCTGGAGGCCCACGACGCGCAAGGGGATGTTCCAGTCACTGTTACTGTCCAC 229

Qy         50 GluHis-----------------CysProSerGlnValThrValLysAlaGluLeuLeu 63
              :::                             |||:::   ||||||
Db        230 GACTTCCCAGGCAAAAAACTAGTGCTGTCCAGTGAGAAGACTGTGCTGACCCCTGCCACC 289

Qy         64 LysThrAlaSerAsnLeuThrValSerVal---LeuGluAlaGluGlyValPheGluLys 82
                          |||:::|||   ::::::               |||         ||||||
Db        290 AACCACATGGGCAACGTCACCTTCACGATCCCAGCCAACAGGGAGTTCAAGTCAGAAAAG 349

Qy         83 GlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyr 102
              |||      |||   :::|||:::   :::                     ::::::
Db        350 GGGCGCAACAAGTTCGTGACCGTGCAGGCC---------ACCTTCGGGACCCAAGTGGTG 400

Qy        103 GluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeu 122
              |||      |||            |||      |||
Db        401 GAGAAGGTGGTGCTGGTCAGCCTGCAGAGCGGGTACCTC-------------------- 439

Qy        123 SerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysPro 142
                                   ||||||||||||||||||   :::|||   |||
Db        440 --------------------------TTCATCCAGACAGACAAGACCATCTACACCCCT 472

Qy        143 LysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThr 162
                          |||   :::||||||   |||:::               |||
Db        473 GGCTCCACAGTTCTCTATCGGATCTTCACCGTCAACCACAAGCTGCTACCCGTGGGCCGG 532

Qy        163 SerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGln---TrpLeuSer 181
              ::::::   :::   |||:::::::|||:::            ::::::|||      ||||||
Db        533 ACGGTCATGGTCAACATTGAGAACCCGGAAGGCATCCCGGTCAAGCAGGACTCCTTGTCT 592

Qy        182 GlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeu 201
                 |||:::   |||||||||               ::::::   :::            :::
Db        593 TCTCAGAACCAGCTTGGCGTCTTGCCCTTGTCTTGGGACATTCCGGAACTCGTCAACATG 652

Qy        202 GlyAspTrpSerIleGlnValGlnValAsnAsp---------GlnThrTyrTyrGlnSer 218
              |||   |||   |||:::                  :::         |||   :::
Db        653 GGCCAGTGGAAGATCCGAGCCTACTATGAAAACTCACCACAGCAGGTCTTCTCCACTGAG 712

Qy        219 PheGlnValSerGluTyrValLeuProLysPheGluValThrLeuGln------ThrPro 236
              |||:::|||   |||||||||||||||   |||||||||      ::::::
Db        713 TTTGAGGTGAAGGAGTACGTGCTGCCCAGTTTCGAGGTCATAGTGGAGCCTACAGAGAAA 772

Qy        237 LeuTyrCysSerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyr 256
                 |||            |||   |||   |||      ||||||||||||::::::   |||
Db        773 TTCTACTACATCTATAACGAGAAGGGCCTGGAGGTCACCATCACCGCCAGGTTCCTCTAC 832

Qy        257 GlyLysProValLysGlyAspValThrLeuThrPhe------------------------ 268
              ||||||   |||:::|||            :::   |||
Db        833 GGGAAGAAAGTGGAGGGAACTGCCTTTGTCATCTTCGGGATCCAGGATGGCGAACAGAGG 892

Qy        269 --------------------------LeuProLeuSerPheTrpGlyLysLysLysAsn 279
                                        :::|||:::                  :::
Db        893 ATTTCCCTGCCTGAATCCCTCAAGCGCATTCCGATTGAGGATGGCTCGGGGGAGGTTGTG 952

Qy        280 IleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMet 299
              :::::::::         ::::::|||      |||               |||
Db        953 CTGAGCCGGAAGGTACTGCTGGACGGGGTGCAGAACCCCCGAGCAGAAGACCTGGTGGGG 1012
```

```
Qy    300 LysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGly 319
Db   1013 AAGTCTTTGTACGTGTCTGCC--------------------------------------- 1033

Qy    320 ProValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSer 339
Db   1034 ---------------ACCGTCATCTTGCACTCAGGCAGTGACATGGTGCAGGCAGAGCGC 1078

Qy    340 ThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrVal 359
Db   1079 AGCGGGATCCCCATCGTGACCTCTCCCTACCAGATCCACTTCACCAAGACACCCAAGTAC 1138

Qy    360 LeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn--- 378
Db   1139 TTCAAACCAGGAATGCCCTTTGACCTCATGGTGTTCGTGACGAACCCTGATGGCTCTCCA 1198

Qy    379 ------GlnLeuThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArg 396
Db   1199 GCCTACCGAGTCCCCGTGGCAGTCCAGGGCGAGGACACTGTGCAGTCTCTAACCCAGGGA 1258

Qy    397 AsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGln 416
Db   1259 GATGGCGTGGCCAAACTCAGCATCAACACACACCCCAGCCAGAAG--------------- 1303

Qy    417 LysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGlu 436
Db   1304 ------------------CCCTTGAGCATCACGGTGCGCACGAAGAAGCAGGAGCTCTCG 1345

Qy    437 AspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaVal 456
Db   1346 GAGGCAGAGCAG------------------------GCTACCAGGACCATGCAGGCTCTG 1381

Qy    457 HisSerLeuPheLysSerProSerLysThrTyrIleGlnLeuLysThr------------ 472
Db   1382 CCCTACAGCACCGTGGGCAACTCCAACAATTACCTGCATCTCTCAGTGCTACGTACAGAG 1441

Qy    473 ------ArgAspGluAsnIleLysValGlySerProPheGluLeuValValSerGlyAsn 490
Db   1442 CTCAGACCCGGGGAGACCCTCAACGTCAACTTCCTCCTGCGAATGGACCGCGCCCACGAG 1501

Qy    491 LysArgLeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGly 510
Db   1502 GCCAAGATCCGCTACTACACCTACCTGATCATGAACAAGGGCAGGCTGTTGAAGGCGGGA 1561

Qy    511 Lys------------------GlnAsnSerThrMetPheSerLeuThrProGluAsnSer 524
Db   1562 CGCCAGGTGCGAGAGCCCGGCCAGGACCTGGTGGTGCTGCCCCTGTCCATCACCACCGAC 1621

Qy    525 TrpThrProLysAlaCysValIleValTyrTyr------------IleGluAspAspGly 540
Db   1622 TTCATCCCTTCCTTCCGCCTGGTGGCGTACTACACGCTGATCGGTGCCAGCGGCCAGAGG 1681

Qy    541 GluIleIleSerAspValLeuLysIleProValGln------------------LeuVal 554
Db   1682 GAGGTGGTGGCCGACTCCGTGTGGGTGGACGTCAAGGACTCCTGCGTGGGCTCGCTGGTG 1741

Qy    555 PheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGlu--------------- 569
Db   1742 GTAAAAAGCGGC------------------------CAGTCAGAAGACCGGCAGCCTGTA 1777

Qy    570 ProSerGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIle 589
Db   1778 CCTGGGCAGCAGATGACCCTGAAGATAGAGGGTGACCACGGGGCCCGGGTGGTA---CTG 1834

Qy    590 ValAlaValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsn 609
Db   1835 GTGGCCGTGGACAAGGGCGTGTTCGTGCTGAATAAGAAGAACAAACTGACGCAGAGTAAG 1894

Qy    610 ValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSer 629
Db   1895 ATCTGGGACGTGGTGGAGAAGGCAGACATCGGCTGCACCCCGGGCAGTGGGAAGGATTAC 1954

Qy    630 PheAlaValPheGlnGluCysGlyLeuTrpValLeuThrAspAla--------------- 644
Db   1955 GCCGGTGTCTTCTCCGACGCAGGGCTGACCTTCACGAGCAGCAGTGGCCAGCAGACCGCC 2014

Qy    644 ------------------------------------------------------------ 644
Db   2015 CAGAGGGCAGAACTTCAGTGCCCGCAGCCAGCCGCCCGCCGACGCCGTTCCGTGCAGCTC 2074

Qy    645 ------------------------------------------AsnLeuThrLysAspTyrIle 651
Db   2075 ACGGAGAAGCGAATGGACAAAGTCGGCAAGTACCCCAAGGAGCTGCGCAAGTGCTGCGAG 2134

Qy    652 AspGlyValTyrAspAsn---------------AlaGluTyrAlaGluArgPheMetGlu 666
Db   2135 GACGGCATGCGGGAGAACCCCATGAGGTTCTCGTGCCAGCGCCGGACCCGTTTCATCTCC 2194

Qy    667 GluAsnGlu------------------------------GlyHisIleValAspIleHis 676
Db   2195 CTGGGCGAGGCGTGCAAGAAGGTCTTCCTGGACTGCTGCAACTACATCACAGAGCTGCGG 2254

Qy    677 ------------------------------------------------AspPheSerLeu 680
Db   2255 CGGCAGCACGCGCGGGCCAGCCACCTGGGCCTGGCCAGGAGTAACCTGGATGAGGACATC 2314

Qy    681 GlySerSerProHisVal------ArgLysHisPheProGluThrTrpIleTrp------ 696
Db   2315 ATTGCAGAAGAGAACATCGTTTCCCGAAGTGAGTTCCCAGAGAGCTGGCTGTGGAACGTT 2374

Qy    697 ------------------------------LeuAspThrAsnMetGlyTyrArgIleTyr 706
Db   2375 GAGGACTTGAAAGAGCCACCGAAAAATGGAATCTCTACGAAGCTC--------------- 2419

Qy    707 GlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPheVal 726
Db   2420 ------ATGAATATATTTTTGAAAGACTCCATCACCACGTGGGAGATTCTGGCTGTGAGC 2473

Qy    727 IleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGln 746
Db   2474 ATGTCGGACAAGAAAGGGATC---TGTGTGGCAGACCCCTTCGAGGTCACAGTAATGCAG 2530

Qy    747 ProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeu 766
Db   2531 GACTTCTTCATCGACCTGCGGCTACCCTACTCTGTTGTTCGAAACGAGCAGGTGGAAATC 2590

Qy    767 GluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIleGlu--- 785
Db   2591 CGAGCCGTTCTCTACAATTAC---CGGCAGAACCAAGAGCTCAAGGTGAGGGTGGAACTA 2647

Qy    786 ---LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHis 804
Db   2648 CTCCACAATCCAGCCTTCTGCAGCCTGGCCACCACCAAGAGG------------CGTCAC 2695

Qy    805 GlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgPro 824
Db   2696 CAGCAGACCATAACCATCCCCCCCAAGTCCTCGTTGTCCGTTCCATATGTCATCGTGCCG 2755

Qy    825 ------ThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAlaSer 842
Db   2756 CTAAAGACCGGCCTGCAGGAAGTGGAAGTCAAGGCTGCTGTCTACCATCATTTCATCAGT 2815

Qy    843 AspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyrSerGln 862
Db   2816 GACGGTGTCAGGAAGTCCCTGAAGGTCGTGCCGGAAGGAATCAGAATGAACAAAACTGTG 2875

Qy    863 SerIleLeuLeuAspLeuThrAspAsnArgLeu------GlnSerThrLeuLysThrLeu 880
Db   2876 GCTGTT---CGCACCCTGGATCCAGAACGCCTGGGCCGTGAAGGAGTGCAGAAAGAG--- 2929

Qy    881 SerPheSerPheProProAsnThrVal------------------ThrGlySerGluArg 894
Db   2930 ------GACATCCCACCTGCAGACCTCAGTGACCAAGTCCCGGACACCGAGTCTGAGACC 2983

Qy    895 ---------ValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGly--- 910
```

```
                    :::|||  |||  :::      :::       ::::::::
Db       2984 AGAATTCTCCTGCAAGGGACCCCAGTGGCCCAGATGACAGAGGATGCCGTCGACGCGGAA 3043

Qy        911 ---LeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPhe 929
                 |||     ||||||         |||   ||||||||||||||||||||||
Db       3044 CGGCTGAAGCACCTCATTGTGACCCCCTCGGGCTGCGGGGAACAGAACATGATCGGCATG 3103

Qy        930 AlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeu 949
                 |||  :::       :::  ||||||   :::   :::|||
Db       3104 ACGCCCACGGTCATCGCTGTGCATTACCTGGATGAAACGGAGCAGTGGGAGAAGTTCGGC 3163

Qy        950 LysGluLys---------AlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeu 966
                 ||||||          ||||||    :::::::::||||||    ::::::|||
Db       3164 CTAGAGAAGCGGCAGGGGGCCTTGGAGCTCATCAAGAAGGGGTACACCCAGCAGCTGGCC 3223

Qy        967 TyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThr 986
              :::::::::          :::|||:::||||||         |||      ||||||
Db       3224 TTCAGACAACCCAGCTCTGCCTTTGCGGCCTTCGTGAAACGGGCACCC------AGCACC 3277

Qy        987 TrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAsp 1006
              ||||||:::|||:::|||::::::    |||       |||        |||   ||||||
Db       3278 TGGCTGACCGCCTACGTGGTCAAGGTCTTCTCTCTGGCTGTCAACCTCATCGCCATCGAC 3337

Qy       1007 GlnAsnValLeuHisArgThrTyrThrTrpLeu---LysGlyHisGlnLysSerAsnGly 1025
                 ||||||         ||||||         ||||||      :::|||
Db       3338 TCCCAAGTCCTCTGCGGGGCTGTTAAATGGCTGATCCTGGAGAAGCAGAAGCCCGACGGG 3397

Qy       1026 GluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGly----------- 1041
                 |||  :::          |||||||||  |||:::   ||||||
Db       3398 GTCTTCCAGGAGGATGCGCCCGTGATACACCAAGAAATGATTGGTGGATTACGGAACAAC 3457

Qy       1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061
              |||:::      :::   |||||||||:::::::::   ||||||
Db       3458 AACGAGAAAGACATGGCCCTCACGGCCTTTGTTCTCATCTCGCTG--------------- 3502

Qy       1062 TyrGlnProAsnIleAspVal---GlnGluSerIleHis--------------------- 1073
                 |||        |||:::   :::|||   ::::::
Db       3503 ---CAGGAGGCTAAAGATATTTGCGAGGAGCAGGTCAACAGCCTGCCAGGCAGCATCACT 3559

Qy       1074 -----------PheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeu 1089
                          |||||||||:::   :::         :::   :::||||||:::
Db       3560 AAAGCAGGAGACTTCCTTGAAGCCAACTAC---ATGAACCTACAGAGATCCTACACTGTG 3616

Qy       1090 AlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsn 1109
              |||:::      |||||||||:::   :::|||      |||
Db       3617 GCCATTGCTGGCTATGCTCTGGCCCAGATGGGCAGGCTGAAGGGGCCTCTTCTTAACAAA 3676

Qy       1110 MetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSer 1129
                 ||||||                                             :::
Db       3677 TTTCTGACC---------------------------------------------ACAGCC 3691

Qy       1130 LysLeuSerAspSerTrpGln---------ProArgSerLeuAspIleGluValAlaAla 1146
              |||       :::   |||:::            :::      ::::::|||      :::
Db       3692 AAAGATAAGAACCGCTGGGAGGACCCTGGTAAGCAGCTCTACAACGTGGAGGCCACATCC 3751

Qy       1147 TyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIle---ProIleMet 1165
              ||||||||||||:::         ||||||   :::      :::   :::   |||::::::
Db       3752 TATGCCCTCTTGGCC---CTACTGCAGCTAAAAGACTTTGACTTTGTGCCTCCCGTCGTG 3808

Qy       1166 ArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThr 1185
              ||||||||||:::      ||||||         ||||||:::   |||||||||   |||
Db       3809 CGTTGGCTCAATGAACAGAGATACTACGGTGGTGGCTATGGCTCTACCCAGGCCACCTTC 3868

Qy       1186 ValAlaLeuLysAlaLeuSerGluPheAlaAla---LeuMetAsnThrGluArgThrAsn 1204
              :::      :::||||||:::::::::                :::   :::      |||
Db       3869 ATGGTGTTCCAAGCCTTGGCTCAATACCAAAAGGACGCCCCTGACCACCAGGAACTGAAC 3928

Qy       1205 IleGlnValThrValThrGlyProSerSer-----------------ProSerProLeu 1218
              :::   |||::::::      ||||||                    |||   ||||||
Db       3929 CTTGATGTGTCCCTCCAACTGCCCAGCCG-CAGCTCCAAGATCACCCACCGTATCCACTG 3987

Qy       1219 AlaValValGlnProMetAlaVal--------------------------Asn-IleSe 1229
                 :::   ||||||   |||:::                             ::: :::::

Db       3988 GGAATCTGCCAGCCTCCTGCGATCAGAAGAGACCAAGGAAAATGAGGGTTTCACAGTCAC 4047

Qy       1229 rAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSe 1249
              :|||   |||   |||                     :::      :::|||:::   ||||||
Db       4048 AGCTGAAGGAAAAGGCCAAGGCACCTTGTCGGTGGTGACAATGTACCATGCTAAGGCCAA 4107

Qy       1249 rGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAl 1269
                             :::                              |||||||||   |||
Db       4108 AGATCAACTCACCTGTAATAAA--------------------TTCGACCTCAAGGTCAC 4146

Qy       1269 aVal--------------------LysGluAsnLysAspAspLeuAsnHisValAspLe 1282
               :::                     |||      :::|||      |||   :::   ||
Db       4147 CATAAAACCAGCACCGGAAACAGAAAAGAGGCCTCAGGATGCCAAGAACACTATGATCCT 4206

Qy       1282 uAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAs 1302
               |   :::||||||   :::   |||        :::   |||::::::::::::::::::
Db       4207 TGAGATCTGTACCAGGTACCGGGGAGACCAGGATGCCACTATGTCTATATTGGACATATC 4266

Qy       1302 nLeuLeuSerGlyPheMetValProSerGlu---------------AlaIleSerLeuSe 1317
               ::::::::::||||||         ::::::                |||      :::
Db       4267 CATGATGACTGGCTTTGCTCCAGACACAGATGACCTGAAGCAGCTGGCCAATGGTGTTGA 4326

Qy       1317 rGluThrValLysLysValGluTyr---------------AspHisGlyLysLeuAsnLe 1332
                 :::   |||   |||                       |||            |||   ::
Db       4327 CAGATACATCTCCAAGTATGAGCTGGACAAAGCCTTCTCCGATAGGAACACCCTCATCAT 4386

Qy       1332 uTyrLeuAspSerValAsnGluThrGln---PheCysValAsnIleProAlaValArgAs 1351
              :|||||||||   |||:::   ::::::      |||:::                 :::
Db       4387 CTACCTGGACAAGGTCTCACACTCTGAGGATGACTGTCTAGCTTTCAAAGTTCACCAATA 4446

Qy       1351 nPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProAr 1371
                |||   |||         |||         :::|||   :::      ||||||
Db       4447 CTTTAATGTAGAGCTTATCCAGCCTGGAGCAGTCAAGGTCTACGCCTATTACAACCTGGA 4506

Qy       1371 gArgGlnAlaValArgSerTyrAsnSerGluVal---------LysLeuSerSerCysAs 1388
                               |||   |||:::   |||            ||||||:::
Db       4507 GGAAAGCTGTACCCGGTTCTACCATCCGGAAAAGGAGGATGGAAAGCTGAACAAG----- 4561

Qy       1388 pLeuCysSerAspValGlnGlyCysArgProCysGluAsp 1401
                ||||||   |||   :::   ||||||   |||:::
Db       4562 -CTCTGCCGTGAT---GAACTGTGCCGCTGTGCTGAGGAG 4597
```

RESULT 5
US-09-132-271-2
; Sequence 2, Application US/09132271
; Patent No. 6221657
; GENERAL INFORMATION:
; APPLICANT: Harrison, Richard Alexander
; APPLICANT: Farries, Charles Timothy
; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: United States of America
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/132,271
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,126
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Hollie L.
; REGISTRATION NUMBER: 31,321

```
;     REFERENCE/DOCKET NUMBER:  102286.377
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 526-6000
;      TELEFAX:  (617) 526-5000
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5056 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-132-271-2

Alignment Scores:
Pred. No.:               4.44e-88      Length:         5056
Score:                   931.00        Matches:        381
Percent Similarity:      40.88%        Conservative:   287
Best Local Similarity:   23.32%        Mismatches:     609
Query Match:             12.67%        Indels:         358
DB:                      3             Gaps:           63

US-10-020-095-4 (1-1428) x US-09-132-271-2 (1-5056)

Qy          3 GlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeuAlaVal 22
              ||||||  ||||||              ||||||   :::          ||||||:::
Db         65 GGTCCCAGCCTGCTG------------CTCCTGCTACTAACCCACCTCCCCCTGGCTCTG 112

Qy         23 AlaProGlyProArgPheLeuValThrAlaProGlyIleIleArg-------------- 37
                    |||   :::   :::       |||   |||:::|||
Db        113 GGG---AGTCCCATGTACTCTATCATCACCCCCAACATCTTGCGGCTGGAGAGCGAGGAG 169

Qy         38 -----------------------ProGlyGlyAsnValThrIleGlyValGluLeuLeu 49
                                     |||:::|||   :::   |||   :::
Db        170 ACCATGGTGCTGGAGGCCCACGACGCGCAAGGGGATGTTCCAGTCACTGTTACTGTCCAC 229

Qy         50 GluHis-----------------CysProSerGlnValThrValLysAlaGluLeuLeu 63
              :::                    |||:::   ||||||
Db        230 GACTTCCCAGGCAAAAAACTAGTGCTGTCCAGTGAGAAGACTGTGCTGACCCCTGCCACC 289

Qy         64 LysThrAlaSerAsnLeuThrValSerVal---LeuGluAlaGluGlyValPheGluLys 82
                          |||:::|||   :::::              |||         ||||||
Db        290 AACCACATGGGCAACGTCACCTTCACGATCCCAGCCAACAGGGAGTTCAAGTCAGAAAAG 349

Qy         83 GlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyr 102
              |||         |||   :::|||:::   :::                  ::::::
Db        350 GGGCGCAACAAGTTCGTGACCGTGCAGGCC---------ACCTTCGGGACCCAAGTGGTG 400

Qy        103 GluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeu 122
              |||      |||            |||         |||
Db        401 GAGAAGGTGGTGCTGGTCAGCCTGCAGAGCGGGTACCTC--------------------- 439

Qy        123 SerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysPro 142
                                         ||||||||||||||||   :::|||   |||
Db        440 ---------------------------TTCATCCAGACAGACAAGACCATCTACACCCCT 472

Qy        143 LysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThr 162
                       |||   :::||||||   |||:::                  |||
Db        473 GGCTCCACAGTTCTCTATCGGATCTTCACCGTCAACCACAAGCTGCTACCCGTGGGCCGG 532

Qy        163 SerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGln---TrpLeuSer 181
              ::::::   :::      |||::::::|||:::      ::::::|||      ||||||
Db        533 ACGGTCATGGTCAACATTGAGAACCCGGAAGGCATCCCGGTCAAGCAGGACTCCTTGTCT 592

Qy        182 GlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeu 201
                    |||:::   |||||||||:::      ::::::   :::                  :::
Db        593 TCTCAGAACCAGCTTGGCGTCTTGCCCTTGTCTTGGGACATTCCGGAACTCGTCAACATG 652

Qy        202 GlyAspTrpSerIleGlnValGlnValAsnAsp---------GlnThrTyrTyrGlnSer 218
              |||   |||   |||:::               :::         |||   :::
Db        653 GGCCAGTGGAAGATCCGAGCCTACTATGAAAACTCACCACAGCAGGTCTTCTCCACTGAG 712

Qy        219 PheGlnValSerGluTyrValLeuProLysPheGluValThrLeuGln------ThrPro 236
              |||:::|||   ||||||||||||||   |||||||||   ::::::
Db        713 TTTGAGGTGAAGGAGTACGTGCTGCCCAGTTTCGAGGTCATAGTGGAGCCTACAGAGAAA 772

Qy        237 LeuTyrCysSerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyr 256
                 |||         |||   |||   |||      ||||||||||||::::::   |||
Db        773 TTCTACTACATCTATAACGAGAAGGGCCTGGAGGTCACCATCACCGCCAGGTTCCTCTAC 832

Qy        257 GlyLysProValLysGlyAspValThrLeuThrPhe----------------------- 268
              ||||||   |||:::|||                  :::   |||
Db        833 GGGAAGAAAGTGGAGGGAACTGCCTTTGTCATCTTCGGGATCCAGGATGGCGAACAGAGG 892

Qy        269 --------------------------LeuProLeuSerPheTrpGlyLysLysLysAsn 279
                                        :::|||:::                :::
Db        893 ATTTCCCTGCCTGAATCCCTCAAGCGCATTCCGATTGAGGATGGCTCGGGGGAGGTTGTG 952

Qy        280 IleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMet 299
              :::::::::            ::::::|||      |||            |||
Db        953 CTGAGCCGGAAGGTACTGCTGGACGGGGTGCAGAACCCCCGAGCAGAAGACCTGGTGGGG 1012

Qy        300 LysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGly 319
              |||::::::      |||:::
Db       1013 AAGTCTTTGTACGTGTCTGCC------------------------------------------ 1033

Qy        320 ProValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSer 339
                    |||   :::      |||   :::   :::   :::
Db       1034 ---------------ACCGTCATCTTGCACTCAGGCAGTGACATGGTGCAGGCAGAGCGC 1078

Qy        340 ThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrVal 359
              :::   :::                  |||   |||   |||
Db       1079 AGCGGGATCCCCATCGTGACCTCTCCCTACCAGATCCACTTCACCAAGACACCCAAGTAC 1138

Qy        360 LeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn--- 378
                 ||||||   :::   |||         |||   ||||||         ||||||:::
Db       1139 TTCAAACCAGGAATGCCCTTTGACCTCATGGTGTTCGTGACGAACCCTGATGGCTCTCCA 1198

Qy        379 ------GlnLeuThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArg 396
                    ::::::   :::      :::      :::   |||   :::::|||||
Db       1199 GCCTACCGAGTCCCCGTGGCAGTCCAGGGCGAGGACACTGTGCAGTCTCTAACCCAGGGA 1258

Qy        397 AsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGln 416
              :::               |||   :::         :::||||||
Db       1259 GATGGCGTGGCCAAACTCAGCATCAACACACACCCCAGCCAGAAG--------------- 1303

Qy        417 LysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGlu 436
                          |||   |||   |||   :::   :::               |||
Db       1304 -----------------CCCTTGAGCATCACGGTGCGCACGAAGAAGCAGGAGCTCTCG 1345

Qy        437 AspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaVal 456
              ::::::   :::                     :::::::::         |||:::
Db       1346 GAGGCAGAGCAG-----------------------GCTACCAGGACCATGCAGGCTCTG 1381

Qy        457 HisSerLeuPheLysSerProSerLysThrTyrIleGlnLeuLysThr----------- 472
                                 |||      |||:::   |||
Db       1382 CCCTACAGCACCGTGGGCAACTCCAACAATTACCTGCATCTCTCAGTGCTACGTACAGAG 1441

Qy        473 ------ArgAspGluAsnIleLysValGlySerProPheGluLeuValValSerGlyAsn 490
                          |||   :::   |||                  :::      :::
Db       1442 CTCAGACCCGGGGAGACCCTCAACGTCAACTTCCTCCTGCGAATGGACCGCGCCCACGAG 1501

Qy        491 LysArgLeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGly 510
                 :::::::::      :::|||::::::::::::::|||:::|||:::      |||
Db       1502 GCCAAGATCCGCTACTACACCTACCTGATCATGAACAAGGGCAGGCTGTTGAAGGCGGGA 1561

Qy        511 Lys-----------------GlnAsnSerThrMetPheSerLeuThrProGluAsnSer 524
              :::                 |||:::            :::      |||:::
Db       1562 CGCCAGGTGCGAGAGCCCGGCCAGGACCTGGTGGTGCTGCCCCTGTCCATCACCACCGAC 1621

Qy        525 TrpThrProLysAlaCysValIleValTyrTyr-----------IleGluAspAspGly 540
              :::   |||            ::::::   ||||||
Db       1622 TTCATCCCTTCCTTCCGCCTGGTGGCGTACTACACGCTGATCGGTGCCAGCGGCCAGAGG 1681

Qy        541 GluIleIleSerAspValLeuLysIleProValGln------------------LeuVal 554
              |||:::::::::|||   :::   :::   |||:::                  ||||||
Db       1682 GAGGTGGTGGCCGACTCCGTGTGGGTGGACGTCAAGGACTCCTGCGTGGGCTCGCTGGTG 1741
```

```
Qy    555 PheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGlu--------------- 569
Db   1742 GTAAAAAGCGGC------------------------CAGTCAGAAGACCGGCAGCCTGTA 1777

Qy    570 ProSerGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIle 589
Db   1778 CCTGGGCAGCAGATGACCCTGAAGATAGAGGGTGACCACGGGGCCCGGGTGGTA---CTG 1834

Qy    590 ValAlaValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsn 609
Db   1835 GTGGCCGTGGACAAGGGCGTGTTCGTGCTGAATAAGAAGAACAAACTGACGCAGAGTAAG 1894

Qy    610 ValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSer 629
Db   1895 ATCTGGGACGTGGTGGAGAAGGCAGACATCGGCTGCACCCCGGGCAGTGGGAAGGATTAC 1954

Qy    630 PheAlaValPheGlnGluCysGlyLeuTrpValLeuThrAspAla--------------- 644
Db   1955 GCCGGTGTCTTCTCCGACGCAGGGCTGACCTTCACGAGCAGCAGTGGCCAGCAGACCGCC 2014

Qy    644 ------------------------------------------------------------ 644
Db   2015 CAGAGGGCAGAACTTCAGTGCCCGCAGCCAGCCGCCCGCCGACGCCGTTCCGTGCAGCTC 2074

Qy    645 ---------------------------------------AsnLeuThrLysAspTyrIle 651
Db   2075 ACGGAGAAGCGAATGGACAAAGTCGGCAAGTACCCCAAGGAGCTGCGCAAGTGCTGCGAG 2134

Qy    652 AspGlyValTyrAspAsn---------------AlaGluTyrAlaGluArgPheMetGlu 666
Db   2135 GACGGCATGCGGGAGAACCCCATGAGGTTCTCGTGCCAGCGCCGGACCCGTTTCATCTCC 2194

Qy    667 GluAsnGlu-----------------------------GlyHisIleValAspIleHis 676
Db   2195 CTGGGCGAGGCGTGCAAGAAGGTCTTCCTGGACTGCTGCAACTACATCACAGAGCTGCGG 2254

Qy    677 ------------------------------------------------AspPheSerLeu 680
Db   2255 CGGCAGCACGCGCGGGCCAGCCACCTGGGCCTGGCCAGGAGTAACCTGGATGAGGACATC 2314

Qy    681 GlySerSerProHisVal------ArgLysHisPheProGluThrTrpIleTrp------ 696
Db   2315 ATTGCAGAAGAGAACATCGTTTCCCGAAGTGAGTTCCCAGAGAGCTGGCTGTGGAACGTT 2374

Qy    697 ------------------------------LeuAspThrAsnMetGlyTyrArgIleTyr 706
Db   2375 GAGGACTTGAAAGAGCCACCGAAAAATGGAATCTCTACGAAGCTC--------------- 2419

Qy    707 GlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPheVal 726
Db   2420 ------ATGAATATATTTTTGAAAGACTCCATCACCACGTGGGAGATTCTGGCTGTGAGC 2473

Qy    727 IleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGln 746
Db   2474 ATGTCGGACAAGAAAGGGATC---TGTGTGGCAGACCCCTTCGAGGTCACAGTAATGCAG 2530

Qy    747 ProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeu 766
Db   2531 GACTTCTTCATCGACCTGCGGCTACCCTACTCTGTTGTTCGAAACGAGCAGGTGGAAATC 2590

Qy    767 GluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIleGlu--- 785
Db   2591 CGAGCCGTTCTCTACAATTAC---CGGCAGAACCAAGAGCTCAAGGTGAGGGTGGAACTA 2647

Qy    786 ---LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHis 804
Db   2648 CTCCACAATCCAGCCTTCTGCAGCCTGGCCACCACCAAGAGG------------CGTCAC 2695

Qy    805 GlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgPro 824
Db   2696 CAGCAGACCATAACCATCCCCCCCAAGTCCTCGTTGTCCGTTCCATATGTCATCGTGCCG 2755

Qy    825 ------ThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAlaSer 842
Db   2756 CTAAAGACCGGCCTGCAGGAAGTGGAAGTCAAGGCTGCTGTCTACCATCATTTCATCAGT 2815

Qy    843 AspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyrSerGln 862
Db   2816 GACGGTGTCAGGAAGTCCCTGAAGGTCGTGCCGGAAGGAATCAGAATGAACAAAACTGTG 2875

Qy    863 SerIleLeuLeuAspLeuThrAspAsnArgLeu------GlnSerThrLeuLysThrLeu 880
Db   2876 GCTGTT---CGCACCCTGGATCCAGAACGCCTGGGCCGTGAAGGAGTGCAGAAAGAG--- 2929

Qy    881 SerPheSerPheProProAsnThrVal------------------ThrGlySerGluArg 894
Db   2930 ------GACATCCCACCTGCAGACCTCAGTGACCAAGTCCCGGACACCGAGTCTGAGACC 2983

Qy    895 ---------ValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGly--- 910
Db   2984 AGAATTCTCCTGCAAGGGACCCCAGTGGCCCAGATGACAGAGGATGCCGTCGACGCGGAA 3043

Qy    911 ---LeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPhe 929
Db   3044 CGGCTGAAGCACCTCATTGTGACCCCCTCGGGCTGCGGGGAACAGAACATGATCGGCATG 3103

Qy    930 AlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeu 949
Db   3104 ACGCCCACGGTCATCGCTGTGCATTACCTGGATGAAACGGAGCAGTGGGAGAAGTTCGGC 3163

Qy    950 LysGluLys---------AlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeu 966
Db   3164 CTAGAGAAGCGGCAGGGGGCCTTGGAGCTCATCAAGAAGGGGTACACCCAGCAGCTGGCC 3223

Qy    967 TyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThr 986
Db   3224 TTCAGACAACCCAGCTCTGCCTTTGCGGCCTTCGTGAAACGGGCACCC------AGCACC 3277

Qy    987 TrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAsp 1006
Db   3278 TGGCTGACCGCCTACGTGGTCAAGGTCTTCTCTCTGGCTGTCAACCTCATCGCCATCGAC 3337

Qy   1007 GlnAsnValLeuHisArgThrTyrThrTrpLeu---LysGlyHisGlnLysSerAsnGly 1025
Db   3338 TCCCAAGTCCTCTGCGGGGCTGTTAAATGGCTGATCCTGGAGAAGCAGAAGCCCGACGGG 3397

Qy   1026 GluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGly------------ 1041
Db   3398 GTCTTCCAGGAGGATGCGCCCGTGATACACCAAGAAATGATTGGTGGATTACGGAACAAC 3457

Qy   1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061
Db   3458 AACGAGAAAGACATGGCCCTCACGGCCTTTGTTCTCATCTCGCTG--------------- 3502

Qy   1062 TyrGlnProAsnIleAspVal---GlnGluSerIleHis--------------------- 1073
Db   3503 ---CAGGAGGCTAAAGATATTTGCGAGGAGCAGGTCAACAGCCTGCCAGGCAGCATCACT 3559

Qy   1074 ------------PheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeu 1089
Db   3560 AAAGCAGGAGACTTCCTTGAAGCCAACTAC---ATGAACCTACAGAGATCCTACACTGTG 3616

Qy   1090 AlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsn 1109
Db   3617 GCCATTGCTGGCTATGCTCTGGCCCAGATGGGCAGGCTGAAGGGGCCTCTTCTTAACAAA 3676

Qy   1110 MetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSer 1129
Db   3677 TTTCTGACC------------------------------------------------ACAGCC 3691

Qy   1130 LysLeuSerAspSerTrpGln---------ProArgSerLeuAspIleGluValAlaAla 1146
Db   3692 AAAGATAAGAACCGCTGGGAGGACCCTGGTAAGCAGCTCTACAACGTGGAGGCCACATCC 3751

Qy   1147 TyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIle---ProIleMet 1165
```

```
               ||||||||||||:::      ||||||   :::      :::   :::   |||::::::
Db      3752 TATGCCCTCTTGGCC---CTACTGCAGCTAAAAGACTTTGACTTTGTGCCTCCCGTCGTG 3808

Qy      1166 ArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThr 1185
             |||||||||:::   ||||||          ||||||:::  |||||||||   |||
Db      3809 CGTTGGCTCAATGAACAGAGATACTACGGTGGTGGCTATGGCTCTACCCAGGCCACCTTC 3868

Qy      1186 ValAlaLeuLysAlaLeuSerGluPheAlaAla---LeuMetAsnThrGluArgThrAsn 1204
             :::       :::||||||::::::::                 :::   :::      |||
Db      3869 ATGGTGTTCCAAGCCTTGGCTCAATACCAAAAGGACGCCCCTGACCACCAGGAACTGAAC 3928

Qy      1205 IleGlnValThrValThrGlyProSerSer------------------ProSerProLeu 1218
             :::   |||::::::       ||||||                    |||   ||||||
Db      3929 CTTGATGTGTCCCTCCAACTGCCCAGCCG-CAGCTCCAAGATCACCCACCGTATCCACTG 3987

Qy      1219 AlaValValGlnProMetAlaVal--------------------------Asn-IleSe 1229
                :::   ||||||   |||:::                          ::: :::::
Db      3988 GGAATCTGCCAGCCTCCTGCGATCAGAAGAGACCAAGGAAAATGAGGGTTTCACAGTCAC 4047

Qy      1229 rAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSe 1249
             :|||  |||  |||                    :::      :::|||:::  ||||||
Db      4048 AGCTGAAGGAAAAGGCCAAGGCACCTTGTCGGTGGTGACAATGTACCATGCTAAGGCCAA 4107

Qy      1249 rGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAl 1269
                                     :::                  |||||||||   |||
Db      4108 AGATCAACTCACCTGTAATAAA--------------------TTCGACCTCAAGGTCAC 4146

Qy      1269 aVal--------------------LysGluAsnLysAspAspLeuAsnHisValAspLe 1282
              :::                    |||     :::|||      |||   :::   ||
Db      4147 CATAAAACCAGCACCGGAAACAGAAAAGAGGCCTCAGGATGCCAAGAACACTATGATCCT 4206

Qy      1282 uAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAs 1302
             |   :::||||||   :::   |||          :::   |||::::::::::::::::
Db      4207 TGAGATCTGTACCAGGTACCGGGGAGACCAGGATGCCACTATGTCTATATTGGACATATC 4266

Qy      1302 nLeuLeuSerGlyPheMetValProSerGlu---------------AlaIleSerLeuSe 1317
             ::::::::::||||||          ::::::                |||      :::
Db      4267 CATGATGACTGGCTTTGCTCCAGACACAGATGACCTGAAGCAGCTGGCCAATGGTGTTGA 4326

Qy      1317 rGluThrValLysLysValGluTyr---------------AspHisGlyLysLeuAsnLe 1332
                :::   |||  |||                        |||          |||   ::
Db      4327 CAGATACATCTCCAAGTATGAGCTGGACAAAGCCTTCTCCGATAGGAACACCCTCATCAT 4386

Qy      1332 uTyrLeuAspSerValAsnGluThrGln---PheCysValAsnIleProAlaValArgAs 1351
             :|||||||||   |||:::   ::::::    |||:::                   :::
Db      4387 CTACCTGGACAAGGTCTCACACTCTGAGGATGACTGTCTAGCTTTCAAAGTTCACCAATA 4446

Qy      1351 nPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProAr 1371
              |||  |||          |||      :::|||   :::      ||||||
Db      4447 CTTTAATGTAGAGCTTATCCAGCCTGGAGCAGTCAAGGTCTACGCCTATTACAACCTGGA 4506

Qy      1371 gArgGlnAlaValArgSerTyrAsnSerGluVal---------LysLeuSerSerCysAs 1388
                           |||   |||:::   |||               ||||||:::
Db      4507 GGAAAGCTGTACCCGGTTCTACCATCCGGAAAAGGAGGATGGAAAGCTGAACAAG----- 4561

Qy      1388 pLeuCysSerAspValGlnGlyCysArgProCysGluAsp 1401
              ||||||   |||   :::   ||||||      |||:::
Db      4562 -CTCTGCCGTGAT---GAACTGTGCCGCTGTGCTGAGGAG 4597
```

RESULT 6
US-09-142-334-23
; Sequence 23, Application US/09142334
; Patent No. 6268485
; GENERAL INFORMATION:
; APPLICANT: Farries, Timothy C.
; APPLICANT: Harrison, Richard A.
; TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
; FILE REFERENCE: 4-30443/A/IMU/PCT
; CURRENT APPLICATION NUMBER: US/09/142,334
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: PCT/GB97/00603
; EARLIER FILING DATE: 1997-03-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
;   LENGTH: 5067
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-142-334-23

```
Alignment Scores:
Pred. No.:               4.46e-88     Length:         5067
Score:                   931.00       Matches:        381
Percent Similarity:      40.88%       Conservative:   287
Best Local Similarity:   23.32%       Mismatches:     609
Query Match:             12.67%       Indels:         358
DB:                      3            Gaps:           63
```

US-10-020-095-4 (1-1428) x US-09-142-334-23 (1-5067)

```
Qy         3 GlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeuAlaVal 22
              ||||||  ||||||                ||||||   :::          ||||||:::
Db        76 GGTCCCAGCCTGCTG------------CTCCTGCTACTAACCCACCTCCCCCTGGCTCTG 123

Qy        23 AlaProGlyProArgPheLeuValThrAlaProGlyIleIleArg--------------- 37
                      |||   :::   :::      |||   |||:::|||
Db       124 GGG---AGTCCCATGTACTCTATCATCACCCCCAACATCTTGCGGCTGGAGAGCGAGGAG 180

Qy        38 -----------------------ProGlyGlyAsnValThrIleGlyValGluLeuLeu 49
                                    |||:::|||   :::   |||   :::
Db       181 ACCATGGTGCTGGAGGCCCACGACGCGCAAGGGGATGTTCCAGTCACTGTTACTGTCCAC 240

Qy        50 GluHis-----------------CysProSerGlnValThrValLysAlaGluLeuLeu 63
             :::                    |||:::   ||||||
Db       241 GACTTCCCAGGCAAAAAACTAGTGCTGTCCAGTGAGAAGACTGTGCTGACCCCTGCCACC 300

Qy        64 LysThrAlaSerAsnLeuThrValSerVal---LeuGluAlaGluGlyValPheGluLys 82
                      |||:::|||   ::::::                 |||            ||||||
Db       301 AACCACATGGGCAACGTCACCTTCACGATCCCAGCCAACAGGGAGTTCAAGTCAGAAAAG 360

Qy        83 GlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyr 102
              |||      |||   :::|||:::   :::                      ::::::
Db       361 GGGCGCAACAAGTTCGTGACCGTGCAGGCC---------ACCTTCGGGACCCAAGTGGTG 411

Qy       103 GluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeu 122
              |||      |||            |||         |||
Db       412 GAGAAGGTGGTGCTGGTCAGCCTGCAGAGCGGGTACCTC-------------------- 450

Qy       123 SerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysPro 142
                                          ||||||||||||||||||      :::|||   |||
Db       451 --------------------------TTCATCCAGACAGACAAGACCATCTACACCCCT 483

Qy       143 LysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThr 162
                      |||   :::||||||   |||:::                   |||
Db       484 GGCTCCACAGTTCTCTATCGGATCTTCACCGTCAACCACAAGCTGCTACCCGTGGGCCGG 543

Qy       163 SerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGln---TrpLeuSer 181
             ::::::   :::   |||::::::|||:::         ::::::|||      ||||||
Db       544 ACGGTCATGGTCAACATTGAGAACCCGGAAGGCATCCCGGTCAAGCAGGACTCCTTGTCT 603

Qy       182 GlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeu 201
                |||:::   |||||||||:::      ::::::   :::                  :::
Db       604 TCTCAGAACCAGCTTGGCGTCTTGCCCTTGTCTTGGGACATTCCGGAACTCGTCAACATG 663

Qy       202 GlyAspTrpSerIleGlnValGlnValAsnAsp---------GlnThrTyrTyrGlnSer 218
              |||   |||   |||:::                   :::         |||   :::
Db       664 GGCCAGTGGAAGATCCGAGCCTACTATGAAAACTCACCACAGCAGGTCTTCTCCACTGAG 723

Qy       219 PheGlnValSerGluTyrValLeuProLysPheGluValThrLeuGln------ThrPro 236
             |||:::|||   |||||||||||||||   |||||||||      ::::::
Db       724 TTTGAGGTGAAGGAGTACGTGCTGCCCAGTTTCGAGGTCATAGTGGAGCCTACAGAGAAA 783

Qy       237 LeuTyrCysSerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyr 256
                |||         |||   |||   |||      ||||||||||||::::::   |||
Db       784 TTCTACTACATCTATAACGAGAAGGGCCTGGAGGTCACCATCACCGCCAGGTTCCTCTAC 843
```

```
Qy      257 GlyLysProValLysGlyAspValThrLeuThrPhe------------------------ 268
            ||||||   |||:::|||             :::    |||
Db      844 GGGAAGAAAGTGGAGGGAACTGCCTTTGTCATCTTCGGGATCCAGGATGGCGAACAGAGG 903

Qy      269 --------------------------LeuProLeuSerPheTrpGlyLysLysLysAsn 279
                                      :::|||:::          :::
Db      904 ATTTCCCTGCCTGAATCCCTCAAGCGCATTCCGATTGAGGATGGCTCGGGGGAGGTTGTG 963

Qy      280 IleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMet 299
            :::::::::        ::::::|||      |||         |||
Db      964 CTGAGCCGGAAGGTACTGCTGGACGGGGTGCAGAACCCCCGAGCAGAAGACCTGGTGGGG 1023

Qy      300 LysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGly 319
            |||::::::      |||:::
Db     1024 AAGTCTTTGTACGTGTCTGCC--------------------------------------- 1044

Qy      320 ProValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSer 339
                           |||   :::      |||   :::   :::   :::
Db     1045 ---------------ACCGTCATCTTGCACTCAGGCAGTGACATGGTGCAGGCAGAGCGC 1089

Qy      340 ThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrVal 359
            :::   :::                   |||   |||   |||
Db     1090 AGCGGGATCCCCATCGTGACCTCTCCCTACCAGATCCACTTCACCAAGACACCCAAGTAC 1149

Qy      360 LeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn--- 378
               ||||||   :::   |||         |||   ||||||      ||||||:::
Db     1150 TTCAAACCAGGAATGCCCTTTGACCTCATGGTGTTCGTGACGAACCCTGATGGCTCTCCA 1209

Qy      379 ------GlnLeuThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArg 396
                  ::::::   :::      :::      :::   |||   ::::::||||||
Db     1210 GCCTACCGAGTCCCCGTGGCAGTCCAGGGCGAGGACACTGTGCAGTCTCTAACCCAGGGA 1269

Qy      397 AsnTyrThrGluTyrTrpSerGlyAsnSerGlyAsnGlnLysMetGluAlaValGln 416
            :::         |||   :::         ::::||||||
Db     1270 GATGGCGTGGCCAAACTCAGCATCAACACACACCCCAGCCAGAAG--------------- 1314

Qy      417 LysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGlu 436
                              |||   |||   |||   :::   :::         |||
Db     1315 ------------------CCCTTGAGCATCACGGTGCGCACGAAGAAGCAGGAGCTCTCG 1356

Qy      437 AspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaVal 456
            ::::::   :::                        :::::::::      |||:::
Db     1357 GAGGCAGAGCAG------------------------GCTACCAGGACCATGCAGGCTCTG 1392

Qy      457 HisSerLeuPheLysSerProSerLysThrTyrIleGlnLeuLysThr------------ 472
                                 |||      |||:::   |||
Db     1393 CCCTACAGCACCGTGGGCAACTCCAACAATTACCTGCATCTCTCAGTGCTACGTACAGAG 1452

Qy      473 ------ArgAspGluAsnIleLysValGlySerProPheGluLeuValValSerGlyAsn 490
                        |||   :::   |||                  :::      :::
Db     1453 CTCAGACCCGGGGAGACCCTCAACGTCAACTTCCTCCTGCGAATGGACCGCGCCCACGAG 1512

Qy      491 LysArgLeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGly 510
               :::::::::         :::|||:::::::::::::::|||:::|||:::   |||
Db     1513 GCCAAGATCCGCTACTACACCTACCTGATCATGAACAAGGGCAGGCTGTTGAAGGCGGGA 1572

Qy      511 Lys------------------GlnAsnSerThrMetPheSerLeuThrProGluAsnSer 524
            :::                  |||:::      :::         |||:::
Db     1573 CGCCAGGTGCGAGAGCCCGGCCAGGACCTGGTGGTGCTGCCCCTGTCCATCACCACCGAC 1632

Qy      525 TrpThrProLysAlaCysValIleValTyrTyr------------IleGluAspAspGly 540
            :::   |||         ::::::   ||||||
Db     1633 TTCATCCCTTCCTTCCGCCTGGTGGCGTACTACACGCTGATCGGTGCCAGCGGCCAGAGG 1692

Qy      541 GluIleIleSerAspValLeuLysIleProValGln------------------LeuVal 554
            |||:::::::::|||   :::   :::   |||:::                  ||||||
Db     1693 GAGGTGGTGGCCGACTCCGTGTGGGTGGACGTCAAGGACTCCTGCGTGGGCTCGCTGGTG 1752

Qy      555 PheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGlu--------------- 569
               |||:::                              ::::::|||
Db     1753 GTAAAAAGCGGC------------------------CAGTCAGAAGACCGGCAGCCTGTA 1788

Qy      570 ProSerGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIle 589
            |||   ::::::::::::|||:::|||                     :::|||   :::
Db     1789 CCTGGGCAGCAGATGACCCTGAAGATAGAGGGTGACCACGGGGCCCGGGTGGTA---CTG 1845

Qy      590 ValAlaValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsn 609
            |||||||||||||||   |||   ::::::|||      |||   :::|||
Db     1846 GTGGCCGTGGACAAGGGCGTGTTCGTGCTGAATAAGAAGAACAAACTGACGCAGAGTAAG 1905

Qy      610 ValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSer 629
            :::         :::|||   :::   |||      |||         :::
Db     1906 ATCTGGGACGTGGTGGAGAAGGCAGACATCGGCTGCACCCCGGGCAGTGGGAAGGATTAC 1965

Qy      630 PheAlaValPheGlnGluCysGlyLeuTrpValLeuThrAspAla--------------- 644
               ||||||   :::   ||||||         :::   :::
Db     1966 GCCGGTGTCTTCTCCGACGCAGGGCTGACCTTCACGAGCAGCAGTGGCCAGCAGACCGCC 2025

Qy      644 ------------------------------------------------------------ 644

Db     2026 CAGAGGGCAGAACTTCAGTGCCCGCAGCCAGCCGCCCGCCGACGCCGTTCCGTGCAGCTC 2085

Qy      645 ---------------------------------------AsnLeuThrLysAspTyrIle 651
                                                      |||   |||
Db     2086 ACGGAGAAGCGAATGGACAAAGTCGGCAAGTACCCCAAGGAGCTGCGCAAGTGCTGCGAG 2145

Qy      652 AspGlyValTyrAspAsn---------------AlaGluTyrAlaGluArgPheMetGlu 666
            ||||||:::   :::|||               :::         ||||||:::
Db     2146 GACGGCATGCGGGAGAACCCCATGAGGTTCTCGTGCCAGCGCCGGACCCGTTTCATCTCC 2205

Qy      667 GluAsnGlu-----------------------------GlyHisIleValAspIleHis 676
                  |||                                :::|||   ::::::
Db     2206 CTGGGCGAGGCGTGCAAGAAGGTCTTCCTGGACTGCTGCAACTACATCACAGAGCTGCGG 2265

Qy      677 ------------------------------------------------AspPheSerLeu 680
                                                            |||   :::
Db     2266 CGGCAGCACGCGCGGGCCAGCCACCTGGGCCTGGCCAGGAGTAACCTGGATGAGGACATC 2325

Qy      681 GlySerSerProHisVal------ArgLysHisPheProGluThrTrpIleTrp------ 696
               :::   ::::::            |||      |||||||||:::|||:::|||
Db     2326 ATTGCAGAAGAGAACATCGTTTCCCGAAGTGAGTTCCCAGAGAGCTGGCTGTGGAACGTT 2385

Qy      697 ------------------------------LeuAspThrAsnMetGlyTyrArgIleTyr 706
                                          :::   |||   :::
Db     2386 GAGGACTTGAAAGAGCCACCGAAAAATGGAATCTCTACGAAGCTC--------------- 2430

Qy      707 GlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPheVal 726
                  :::      :::   |||||||||||||:::|||
Db     2431 ------ATGAATATATTTTGAAAGACTCCATCACCACGTGGGAGATTCTGGCTGTGAGC 2484

Qy      727 IleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGln 746
            :::|||:::         |||:::                  |||   |||:::   |||
Db     2485 ATGTCGGACAAGAAAGGGATC---TGTGTGGCAGACCCCTTCGAGGTCACAGTAATGCAG 2541

Qy      747 ProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeu 766
               |||||||||   |||   |||||||||||||||:::|||   |||:::      :::
Db     2542 GACTTCTTCATCGACCTGCGGCTACCCTACTCTGTTGTTCGAAACGAGCAGGTGGAAATC 2601

Qy      767 GluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIleGlu--- 785
                  ::::::||||||   :::            |||:::||||||   :::|||
Db     2602 CGAGCCGTTCTCTACAATTAC---CGGCAGAACCAAGAGCTCAAGGTGAGGGTGGAACTA 2658

Qy      786 ---LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHis 804
               :::      |||      |||   |||:::                     |||
Db     2659 CTCCACAATCCAGCCTTCTGCAGCCTGGCCACCACCAAGAGG------------CGTCAC 2706

Qy      805 GlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgPro 824
            |||||||||:::   :::|||   :::         :::|||   :::   |||   |||
Db     2707 CAGCAGACCATAACCATCCCCCCCAAGTCCTCGTTGTCCGTTCCATATGTCATCGTGCCG 2766

Qy      825 ------ThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAlaSer 842
                  |||   |||   |||:::   :::                        |||
Db     2767 CTAAAGACCGGCCTGCAGGAAGTGGAAGTCAAGGCTGCTGTCTACCATCATTTCATCAGT 2826

Qy      843 AspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyrSerGln 862
            |||   |||   :::   :::   |||      ||||||||         :::   :::
```

```
Db      2827 GACGGTGTCAGGAAGTCCCTGAAGGTCGTGCCGGAAGGAATCAGAATGAACAAAACTGTG 2886

Qy       863 SerIleLeuLeuAspLeuThrAspAsnArgLeu------GlnSerThrLeuLysThrLeu 880
             ::::::          |||         ||||||       :::        |||
Db      2887 GCTGTT---CGCACCCTGGATCCAGAACGCCTGGGCCGTGAAGGAGTGCAGAAAGAG--- 2940

Qy       881 SerPheSerPheProProAsnThrVal------------------ThrGlySerGluArg 894
                         ||||||      :::                   |||   ||||||
Db      2941 ------GACATCCCACCTGCAGACCTCAGTGACCAAGTCCCGGACACCGAGTCTGAGACC 2994

Qy       895 ---------ValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGly--- 910
                      :::|||   |||    :::       :::        :::::::::
Db      2995 AGAATTCTCCTGCAAGGGACCCCAGTGGCCCAGATGACAGAGGATGCCGTCGACGCGGAA 3054

Qy       911 ---LeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPhe 929
                |||      ||||||      |||   ||||||||||||||||||||||
Db      3055 CGGCTGAAGCACCTCATTGTGACCCCCTCGGGCTGCGGGGAACAGAACATGATCGGCATG 3114

Qy       930 AlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeu 949
                |||   :::      :::   ||||||   :::   :::|||
Db      3115 ACGCCCACGGTCATCGCTGTGCATTACCTGGATGAAACGGAGCAGTGGGAGAAGTTCGGC 3174

Qy       950 LysGluLys---------AlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeu 966
                ||||||         ||||||      :::::::::||||||   ::::::|||
Db      3175 CTAGAGAAGCGGCAGGGGGCCTTGGAGCTCATCAAGAAGGGGTACACCCAGCAGCTGGCC 3234

Qy       967 TyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThr 986
             :::::::::          :::|||:::||||||               |||   ||||||
Db      3235 TTCAGACAACCCAGCTCTGCCTTTGCGGCCTTCGTGAAACGGGCACCC------AGCACC 3288

Qy       987 TrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAsp 1006
             ||||||:::|||:::|||::::::      |||         |||         |||   ||||||
Db      3289 TGGCTGACCGCCTACGTGGTCAAGGTCTTCTCTCTGGCTGTCAACCTCATCGCCATCGAC 3348

Qy      1007 GlnAsnValLeuHisArgThrTyrThrTrpLeu---LysGlyHisGlnLysSerAsnGly 1025
                      ||||||            ||||||         ||||||   :::|||
Db      3349 TCCCAAGTCCTCTGCGGGGCTGTTAAATGGCTGATCCTGGAGAAGCAGAAGCCCGACGGG 3408

Qy      1026 GluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGly------------ 1041
                |||   :::         |||||||||   |||:::   ||||||
Db      3409 GTCTTCCAGGAGGATGCGCCCGTGATACACCAAGAAATGATTGGTGGATTACGGAACAAC 3468

Qy      1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061
             |||:::         :::   |||||||||:::::::::   ||||||
Db      3469 AACGAGAAAGACATGGCCCTCACGGCCTTTGTTCTCATCTCGCTG--------------- 3513

Qy      1062 TyrGlnProAsnIleAspVal---GlnGluSerIleHis--------------------- 1073
                   |||      |||:::   :::|||   ::::::
Db      3514 ---CAGGAGGCTAAAGATATTTGCGAGGAGCAGGTCAACAGCCTGCCAGGCAGCATCACT 3570

Qy      1074 ------------PheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeu 1089
                         |||||||||:::   :::         :::      :::||||||:::
Db      3571 AAAGCAGGAGACTTCCTTGAAGCCAACTAC---ATGAACCTACAGAGATCCTACACTGTG 3627

Qy      1090 AlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsn 1109
             |||:::      |||||||||:::   :::|||         |||
Db      3628 GCCATTGCTGGCTATGCTCTGGCCCAGATGGGCAGGCTGAAGGGGCCTCTTCTTAACAAA 3687

Qy      1110 MetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSer 1129
                ||||||                                                 :::
Db      3688 TTTCTGACC------------------------------------------ACAGCC 3702

Qy      1130 LysLeuSerAspSerTrpGln---------ProArgSerLeuAspIleGluValAlaAla 1146
             |||   :::      |||:::                :::      ::::::|||      :::
Db      3703 AAAGATAAGAACCGCTGGGAGGACCCTGGTAAGCAGCTCTACAACGTGGAGGCCACATCC 3762

Qy      1147 TyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIle---ProIleMet 1165
             ||||||||||||:::         ||||||   :::      :::   :::   |||::::::
Db      3763 TATGCCCTCTTGGCC---CTACTGCAGCTAAAAGACTTTGACTTTGTGCCTCCCGTCGTG 3819

Qy      1166 ArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThr 1185
             |||||||||:::   ||||||            ||||||:::   |||||||||   |||
Db      3820 CGTTGGCTCAATGAACAGAGATACTACGGTGGTGGCTATGGCTCTACCCAGGCCACCTTC 3879

Qy      1186 ValAlaLeuLysAlaLeuSerGluPheAlaAla---LeuMetAsnThrGluArgThrAsn 1204
             :::      :::||||||:::::::::                :::   :::         |||
Db      3880 ATGGTGTTCCAAGCCTTGGCTCAATACCAAAAGGACGCCCCTGACCACCAGGAACTGAAC 3939

Qy      1205 IleGlnValThrValThrGlyProSerSer------------------ProSerProLeu 1218
             :::   |||::::::         ||||||                   |||   ||||||
Db      3940 CTTGATGTGTCCCTCCAACTGCCCAGCCG-CAGCTCCAAGATCACCCACCGTATCCACTG 3998

Qy      1219 AlaValValGlnProMetAlaVal--------------------------Asn-IleSe 1229
                :::   ||||||   |||:::                          ::: :::::
Db      3999 GGAATCTGCCAGCCTCCTGCGATCAGAAGAGACCAAGGAAAATGAGGGTTTCACAGTCAC 4058

Qy      1229 rAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSe 1249
              :|||   |||   |||                  :::      :::|||:::   ||||||
Db      4059 AGCTGAAGGAAAAGGCCAAGGCACCTTGTCGGTGGTGACAATGTACCATGCTAAGGCCAA 4118

Qy      1249 rGlySerSerArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAl 1269
                               :::                        |||||||||   |||
Db      4119 AGATCAACTCACCTGTAATAAA--------------------TTCGACCTCAAGGTCAC 4157

Qy      1269 aVal--------------------LysGluAsnLysAspAspLeuAsnHisValAspLe 1282
              :::                    |||      :::|||      |||   :::   ||
Db      4158 CATAAAACCAGCACCGGAAACAGAAAAGAGGCCTCAGGATGCCAAGAACACTATGATCCT 4217

Qy      1282 uAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAs 1302
             |   :::||||||   :::   |||      :::   |||::::::::::::::::::
Db      4218 TGAGATCTGTACCAGGTACCGGGGAGACCAGGATGCCACTATGTCTATATTGGACATATC 4277

Qy      1302 nLeuLeuSerGlyPheMetValProSerGlu---------------AlaIleSerLeuSe 1317
             :::::::::||||||               ::::::             |||   :::
Db      4278 CATGATGACTGGCTTTGCTCCAGACACAGATGACCTGAAGCAGCTGGCCAATGGTGTTGA 4337

Qy      1317 rGluThrValLysLysValGluTyr---------------AspHisGlyLysLeuAsnLe 1332
                :::   |||   |||                      |||         |||   ::
Db      4338 CAGATACATCTCCAAGTATGAGCTGGACAAAGCCTTCTCCGATAGGAACACCCTCATCAT 4397

Qy      1332 uTyrLeuAspSerValAsnGluThrGln---PheCysValAsnIleProAlaValArgAs 1351
             :||||||||   |||:::   ::::::      |||:::            :::
Db      4398 CTACCTGGACAAGGTCTCACACTCTGAGGATGACTGTCTAGCTTTCAAAGTTCACCAATA 4457

Qy      1351 nPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProAr 1371
               |||   |||         |||      :::|||   :::      ||||||
Db      4458 CTTTAATGTAGAGCTTATCCAGCCTGGAGCAGTCAAGGTCTACGCCTATTACAACCTGGA 4517

Qy      1371 gArgGlnAlaValArgSerTyrAsnSerGluVal---------LysLeuSerSerCysAs 1388
                      |||   |||:::   |||               ||||||:::
Db      4518 GGAAAGCTGTACCCGGTTCTACCATCCGGAAAAGGAGGATGGAAAGCTGAACAAG----- 4572

Qy      1388 pLeuCysSerAspValGlnGlyCysArgProCysGluAsp 1401
               ||||||   |||   :::   ||||||      |||:::
Db      4573 -CTCTGCCGTGAT---GAACTGTGCCGCTGTGCTGAGGAG 4608
```

RESULT 7
US-08-662-227-1
; Sequence 1, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
;   APPLICANT: VOGEL, CARL-WILHELM
;   APPLICANT: BREDEHORST, REINHORST
;   APPLICANT: KOCK, MICHAEL
;   APPLICANT: FRITZINGER, DAVID
;   TITLE OF INVENTION: RECOMBINANT PROCVF
;   NUMBER OF SEQUENCES: 39
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/662,227
;       FILING DATE:  14-JUN-1996
;       CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;       NAME:  OBLON, NORMAN F.
;       REGISTRATION NUMBER:  24,618
;       REFERENCE/DOCKET NUMBER:  1126-0107-0X
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-413-3000
;       TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH:  5948 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-662-227-1

Alignment Scores:
Pred. No.:                5.24e-84      Length:         5948
Score:                    894.00        Matches:        361
Percent Similarity:       40.90%        Conservative:   295
Best Local Similarity:    22.51%        Mismatches:     672
Query Match:              12.17%        Indels:         276
DB:                       2             Gaps:           56

US-10-020-095-4 (1-1428) x US-08-662-227-1 (1-5948)

Qy         18 AlaAlaLeuAlaVal---------------AlaProGlyProArgPheLeuValThrAla 32
              ||||||||   :::                :::   |||      :::   :::
Db         31 GCTGCTCTATTGATTGGTTTTCCAGGGTCTTCTCATGGGGCTCTCTACACCCTCATCACC 90

Qy         33 ProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeu---GluHis 51
              |||   ::::::|||                   |||   ||||||         :::
Db         91 CCTGCTGTTTTGCGAACAGACACAGAAGAGCAAATTTTGGTGGAGGCCCATGGAGACAGT 150

Qy         52 CysProSerGlnValThrValLysAla------------GluLeuLeuLysThrAlaSer 67
                 |||   |||:::   :::                            |||||
Db        151 ACTCCAAAACAGCTTGACATCTTTGTTCATGATTTTCCACGGAAGCAGAAAACCTTGTTC 210

Qy         68 AsnLeuThrValSerValLeuGluAlaGluGlyValPheGluLysGlySerPheLysThr 87
                       |||   :::      |||   |||:::                     |||
Db        211 CAAACCAGAGTAGATATGAATCCAGCAGGAGGCATGCTTGTCACT---------CCAACT 261

Qy         88 LeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyrGluLeuArgValThr 107
              :::   :::|||:::      :::::::::            :::         |||
Db        262 ATAGAGATTCCAGCAAAAGAAGTGAGTACGGACTCCAGGCAAAATCAATATGTGGTTGTG 321

Qy        108 GlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThrLys 127
                    |||      ::::::                     |||||:::::::::
Db        322 CAAGTAACTGGTCCTCAAGTGAGATTGGAAAAGGTGGTTCTCCTTTCTTACCAGAGTAGC 381

Qy        128 ArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluValLys 147
                    :::||||||||||||||||||   :::|||   |||         |||
Db        382 ------TTTCTGTTTATCCAGACAGATAAAGGCATCTATACACCAGGGTCTCCAGTACTC 435

Qy        148 PheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIleLeu 167
              :::|||:::   ::::::      :::               ::::::   :::
Db        436 TATCGTGTTTTTCTATGGATCACAACACAAGCAAGATGAACAAAACTGTGATTGTTGAG 495

Qy        168 IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAspLeuGly 187
                 :::   |||:::      |||:::               |||         ||||||
Db        496 TTTCAGACTCCAGAAGGCATTCTTGTCAGT---------TCTAATTCAGTTGACCTAAAC 546

Qy        188 ValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSerIle--- 206
                       :::   |||               ||||||   |||   |||
Db        547 TTCTTC---TGGCCTTACAATTTACCAGACCTTGTCAGTTTGGGGACTTGGAGGATTGTG 603

Qy        207 ---GlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrVal 225
                 :::   :::   :::   :::   |||            |||   |||   :::||||||
Db        604 GCCAAATATGAACATTCCCCAGAGAATTATACTGCATATTTGATGTCAGGAAATATGTG 663

Qy        226 LeuProLysPheGluValThrLeuGlnThrPro---------LeuTyrCysSerMetAsn 242
                ||||||   |||||||||   ||||||                    |||         |||
Db        664 TTGCCAAGCTTTGAAGTCCGTCTGCAACCATCAGAGAAGTTTTTTACATTGACGGCAAT 723

Qy        243 SerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGly 262
                 ::::::   :::   :::|||||||||:::|||   ||||||:::   |||:::|||
Db        724 ---GAAAATTTCCACGTGTCTATCACTGCAAGGTACTTGTATGGAGAGGAAGTGGAAGGT 780

Qy        263 AspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle------ 280
                 |||               :::                  ||||||:::|||
Db        781 ---GTGGCCTTTGTCCTCTTTGGAGTGAAAATAGATGATGCTAAAAAGAGTATTCCAGAC 837

Qy        281 ------ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGlu 298
                    |||:::         |||:::|||                  :::         :::
Db        838 TCACTCACGAGAATTCCGATTATTGATGGAGATGGAAAGCAACACTAAAAAGAGATACA 897

Qy        299 MetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerPro 318
                 ::::::         :::   |||   |||         :::         :::|||
Db        898 TTCCGTTCTCGATTTCCAAATCTCAATGAGCTTGTTGGGCATACTCTGTATGCATCT--- 954

Qy        319 GlyProValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnVal 338
                             :::|||   :::||||||||   :::   :::
Db        955 ---------------GTAACAGTCATGACAGAATCAGGCAGTGATATGGTAGTGACTGAG 999

Qy        339 SerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThr 358
                 :::   :::                        |||      |||   |||
Db       1000 CAAAGCGGCATTCATATTGTGGCATCTCCCTATCAGATCCACTTCACAAAAACCCCCAAA 1059

Qy        359 ValLeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn 378
                 ||||||   :::   :::         ||||||   ||||||         ||||||:::
Db       1060 TATTTCAAGCCAGGAATGCCATATGAACTGACGGTGTATGTTACCAACCCTGATGGCTCA 1119

Qy        379 GlnLeuThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyr 398
                                                 :::   |||                  :::
Db       1120 CCAGCT-----------------------GCCCATGTGCCAGTGGTATCAGAGGCCTTT 1155

Qy        399 ThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIle 418
                          :::   |||                  :::               |||:::
Db       1156 -----------------CATTCTATGGGAACCACTTTGAGTGATGGGACTGCTAAGCTC 1197

Qy        419 AsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSer 438
                    :::|||   :::                     ||||||            :::
Db       1198 ATCCTGAACATACCATTGAATGCTCAA---------AGCCTACCAATCACTGTTAGAACT 1248

Qy        439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
              :::         |||               :::||||||         |||:::
Db       1249 AACCATGGAGACCTCCCAAGAGAACGCCAGGCAACAAAGTCCATGACAGCCATAGCCTAC 1308

Qy        459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
                 :::         |||      |||:::   :::                  ||||||
Db       1309 CAAACCCAGGGAGGATCTGGAAACTATCTTCATGTAGCCATTACATCTACAGAGATTAAG 1368

Qy        479 ValGlySer------ProPheGluLeuValValSerGlyAsn-----------LysArg 492
                 |||         |||            |||   ||||||              |||:::
Db       1369 CCCGGAGATAACTTACCTGTCAATTTCAATGTGAAGGGCAATGCAAATTCACTGAAGCAG 1428

Qy        493 LeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLys--- 511
              :::|||      :::|||:::::::::::::::|||::::::         ||||||:::
Db       1429 ATCAAATATTTCACATACCTCATATTGAATAAAGGGAAGATTTTCAAGGTTGGCAGGCAA 1488

Qy        512 ---------------GlnAsnSerThrMetPheSerLeuThrProGluAsnSerTrpThr 526
                              ||||||            :::|||
Db       1489 CCCAGGAGAGATGGGCAGAATCTGGTGACCATGAATCTGCATATCACTCCAGATCTCATC 1548

Qy        527 ProLysAlaCysValIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspVal 546
              |||         :::   ||||||            :::   ||||||::::::|||
Db       1549 CCTTCCTTCCGGTTTGTGGCTTACTACCAAGTGGGAAACAACGAAATTGTGGCTGATTCT 1608
```

```
Qy        547 LeuLysIleProValGln------------------LeuValPheLys------AsnLys 558
Db       1609 GTCTGGGTGGATGTGAAGGATACCTGCATGGGAACGTTGGTTGTGAAAGGAGACAATCTA 1668

Qy        559 IleLysLeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIle 578
Db       1669 ATACAAATG------------------------CCAGGAGCTGCAATGAAAATCAAATTG 1704

Qy        579 SerValThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeu 598
Db       1705 GAA---GGGGATCCAGGTGCTCGGGTTGGTCTTGTGGCTGTGGACAAAGCAGTATATGTT 1761

Qy        599 MetAsnAlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsn 618
Db       1762 CTCAATGATAAATATAAGATTAGCCAAGCTAAGATATGGGACACAATAGAAAAGAGTGAC 1821

Qy        619 ThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeu 638
Db       1822 TTTGGCTGTACAGCTGGCAGTGGCCAGAATAATCTGGGTGTGTTTGAAGATGCTGGACTG 1881

Qy        639 TrpValLeuThrAspAlaAsnLeu---ThrLysAsp------------------------ 649
Db       1882 GCTCTGACAACCAGCACTAATCTCAACACCAAACAGAGATCAGCTGCAAAGTGTCCTCAG 1941

Qy        649 ------------------------------------------------------------ 649
Db       1942 CCTGCAAATCGGAGGCGTCGCAGTTCTGTTTTGCTGCTTGACAGCAACGCAAGCAAAGCG 2001

Qy        649 ------------------------------------------------------------ 649
Db       2002 GCAGAATTTCAGGATCAAGACCTGCGTAAATGCTGTGAAGATGTCATGCATGAGAACCCC 2061

Qy        649 ------------------------------------------------------------ 649
Db       2062 ATGGGGTACACTTGTGAAAAGCGTGCAAAATACATCCAGGAGGGAGATGCTTGTAAGGCT 2121

Qy        650 ---------------------TyrIleAspGlyValTyrAsp---AsnAlaGluTyrAla 661
Db       2122 GCCTTCCTTGAATGCTGTCGCTACATCAAGGGGGTCCGAGATGAAAACCAACGGGAGAGC 2181

Qy        662 GluArgPheMet------GluGluAsnGlu---GlyHisIleValAspIleHisAspPhe 678
Db       2182 GAGTTGTTTCTGGCAAGAGATGATAATGAAGATGGTTTCATAGCAGATAGTGATATCATC 2241

Qy        679 SerLeuGlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeu--- 697
Db       2242 TCA---------------------AGGTCTGATTTCCCCAAGAGTTGGTTGTGGCTAACA 2280

Qy        698 ---------------AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThr 712
Db       2281 AAGGACTTGACCGAGGAGCCTAACAGTCAAGGGATTTCAAGCAAGACAATGTCTTTTTAT 2340

Qy        713 ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732
Db       2341 CTGAGGGATTCCATCACAACCTGGGTGGTGCTGGCTGTAAGCTTTACACCCACCAAAGGG 2400

Qy        733 LeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
Db       2401 ATC---TGTGTGGCTGAACCTTATGAAATAAGAGTCATGAAAGTCTTCTTCATTGATCTT 2457

Qy        753 AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772
Db       2458 CAAATGCCATATTCAGTAGTGAAGAATGAGCAGGTGGAGATTCGAGCTATTCTGCACAAC 2517

Qy        773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
Db       2518 TACGTTAACGAGGATATTTATGTGCGAGTGGAACTGTTATACAACCCAGCCTTC------ 2571

Qy        793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db       2572 ------TGCAGTGCTTCCACAAAAGGACAAAGATACCGACAGCAGTTCCCAATTAAAGCC 2625

Qy        813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly------GluIle 830
Db       2626 CTGTCCTCCAGAGCAGTACCGTTTGTGATAGTCCCATTAGAGCAAGGATTGCATGATGTT 2685

Qy        831 ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeu 850
Db       2686 GAGATTAAAGCAAGTGTCCAGGAAGCGTTGTGGTCAGACGGTGTGAGGAAGAAACTGAAA 2745

Qy        851 ValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870
Db       2746 GTTGTACCTGAAGGGGTACAGAAATCCATTGTGACTATTGTTAAACTGGACCCAAGGGCA 2805

Qy        871 AsnArgLeuGlnSerThr------------------LeuLysThrLeuSerPheSerPhe 884
Db       2806 AAAGGAGTTGGTGGAACACAGCTAGAAGTGATCAAAGCCCGCAAATTAGATGACAGAGTG 2865

Qy        885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904
Db       2866 CCTGACACAGAAATTGAAACCAAGATTATCATCCAAGGTGACCCTGTGGCTCAGATTATT 2925

Qy        905 GlyProSerIleAsnGly------LeuAlaSerLeuIleArgMetProTyrGlyCysGly 922
Db       2926 GAAAACTCAATTGATGGAAGTAAACTCAACCATCTCATTATCACTCCTTCTGGCTGTGGG 2985

Qy        923 GluGlnAsnMetIleAsnPheAlaProAsnIle---------TyrIleLeuAspTyrLeu 939
Db       2986 GAGCAAAATATGATCCGCATGGCCGCACCAGTTATTGCCACCTACTACCTGGACACCACA 3045

Qy        940 ThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGln 959
Db       3046 GAGCAGTGGGAGACTCTCGGCATAAATCGCAGGACTGAAGCTGTCAATCAGATCGTGACT 3105

Qy        960 GlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979
Db       3106 GGTTATGCCCAGCAGATGGTGTACAAGAAAGCAGATCATTCCTATGCAGCATTTACAAAC 3165

Qy        980 TyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPhe---LeuGlu 998
Db       3166 ------CGTGCATCTAGTTCTTGGCTAACAGCATATGTCGTAAAAGTCTTTGCCATGGCT 3219

Qy        999 AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeu--- 1017
Db       3220 GCCAAAATGGTAGCAGGCATTAGTCATGAAATCATTTGTGGAGGTGTGAGGTGGCTGATT 3279

Qy       1018 LysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGlu 1037
Db       3280 CTGAACAGGCAACAACCAGATGGAGCGTTCAAAGAAAATGCCCCTGTACTTTCTGGAACA 3339

Qy       1038 LeuGlnGly---------GlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db       3340 ATGCAGGGAGGAATTCAAGGTGCTGAAGAAGAAGTATATTTAACAGCTTTCATTCTGGTT 3399

Qy       1055 SerLeuLeuGlyTyrArg------------LysTyrGlnProAsnIleAsp------Val 1068
Db       3400 GCGTTGTTGGAATCCAAAACAATCTGCAATGACTATGTCAATAGTCTAGACAGCAGCATC 3459

Qy       1069 GlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db       3460 AAGAAGGCCACAAATTATTTACTCAAAAAGTATGAGAAA---CTGCAAAGGCCTTACACT 3516

Qy       1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeu 1108
Db       3517 ACAGCCCTCACAGCCTATGCTTTGGCT------------------GCTGCAGACCAACTC 3558

Qy       1109 AsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGlu 1128
Db       3559 AATGATGACAGGGTACTCATGGCAGCATCAACAGGAAGGGATCATTGG------------ 3606

Qy       1129 SerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAla 1148
Db       3607 ------------GAAGAATACAATGCTCACACCCACAACATTGAAGGCACTTCCTATGCC 3654

Qy       1149 LeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeu 1168
```

```
Db      3655 TTGTTGGCCCTGCTGAAAATGAAGAAATTTGATCAAACTGGTCCCATAGTCAGATGGCTG 3714

Qy      1169 SerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeu 1188
              :::    |||              :::    ||||||   |||   :::|||
Db      3715 ACAGATCAGAATTTTTATGGGGAAACATATGGACAAACCCAAGCAACAGTTATGGCATTT 3774

Qy      1189 LysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg---ThrAsnIleGlnVal 1207
             :::||||||:::|||:::          |||   |||    :::      |||:::   :::
Db      3775 CAAGCTCTTGCTGAATATGAGATTCAGATGCCTACCCATAAGGACTTAAACTTAGATATT 3834

Qy      1208 ThrValThrGlyProSerSerProSerProLeu-------------------------- 1218
             |||:::      |||      |||:::
Db      3835 ACTATTGAACTGCCAGATCGAGAAGTACCTATAAGGTACAGAATTAATTATGAAAATGCT 3894

Qy      1219 ---------AlaValValGlnPro---------MetAlaValAsnIleSerAlaAsnGly 1232
                            |||:::              :::   ::::::|||:::|||
Db      3895 CTCCTGGCTCGGACAGTAGAGACCAAACTCAACCAAGACATCACTGTGACAGCATCAGGT 3954

Qy      1233 PheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSer 1252
                |||   |||           :::       ||||||   :::
Db      3955 GATGGAAAAGCAACAATGACCATTTTGACATTCTATAACGCACAGTTG----------- 4002

Qy      1253 ArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGlu 1272
             :::   :::   ::::::   |||:::          |||   |||:::|||:::|||:::
Db      4003 CAGGAGAAGGCAAATGTTTGCAATAAA------TTTCATCTTAATGTTTCTGTTGAA--- 4053

Qy      1273 AsnLysAspAspLeuAsnHisValAsp-----------------LeuAsnValCysThr 1286
                      ||||||   :::                        |||   :::||||||
Db      4054 ---AACATCCACTTGAATGCAATGGGAGCCAAGGGAGCCCTCATGCTCAAGATCTGCACA 4110

Qy      1287 SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly 1306
                :::   |||              |||   |||   :::::::::::::::|||:::|||
Db      4111 AGGTATCTGGGAGAAGTTGATTCTACAATGACAATAATTGATATTTCTATGCTGACTGGT 4170

Qy      1307 PheMetValProSerGluAlaIleSer---------------LeuSerGluThrValLys 1321
             |||:::      :::|||   ::::::                :::         :::
Db      4171 TTTCTCCCTGATGCTGAAGACCTTACAAGGCTTTCTAAAGGAGTGGACAGATACATCTCC 4230

Qy      1322 LysValGluTyrAspHisGlyLysLeuAsn---------------LeuTyrLeuAspSer 1336
             :::   |||   |||:::                          :::||||||:::
Db      4231 AGATATGAAGTTGACAATAATATGGCTCAGAAAGTAGCTGTTATCATTTACTTAAACAAG 4290

Qy      1337 ValAsnGluThrGln---PheCysValAsnIleProAlaValArgAsnPheLysValSer 1355
             |||:::   ::::::       |||::::::             :::::::::|||:::|||
Db      4291 GTCTCCCACTCTGAAGATGAATGCCTGCACTTTAAGATTCTCAAGCATTTTGAAGTTGGC 4350

Qy      1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaVal 1375
                   |||      ||||||      :::      ||||||          :::
Db      4351 TTCATTCAGCCAGGATCAGTCAAGGTGTACAGCTACTACAATCTAGATGAAAAATGTACC 4410

Qy      1376 ArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGly 1395
             :::    |||:::   :::                 :::
Db      4411 AAGTTCTACCATCCAGATAAAGGAACAGGCCTTCTCAATAAGATATGTATTGGTAACGTT 4470

Qy      1396 CysArgProCysGluAspGlyAlaSerGlySerHisHis-------------------- 1408
             ||||||         :::      |||      :::|||
Db      4471 TGCCGATGTGCAGGAGAAACCTGTTCCTCGCTCAACCATCAGGAAAGGATTGATGTTCCA 4530

Qy      1409 --------------------HisSerSerValIlePheIlePheCysPheLysLeuLeu 1421
                                 ::::::|||   :::::::::         |||||||||
Db      4531 TTACAAATTGAAAAAGCCTGCGAGACGAATGTGGATTATGTCTACAAAACCAAGCTGCTT 4590

Qy      1422 TyrPheMetGlu 1425
                      |||
Db      4591 CGAATAGAAGAA 4602
```

RESULT 8
US-09-017-947-1
; Sequence 1, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
;   APPLICANT: VOGEL, CARL-WILHELM
;   APPLICANT: BREDEHORST, REINHORST
;   APPLICANT: KOCK, MICHAEL
;   APPLICANT: FRITZINGER, DAVID
;   TITLE OF INVENTION: RECOMBINANT PROCVF
;   NUMBER OF SEQUENCES: 39
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/017,947
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/662,227
;   FILING DATE: 14-JUN-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 1126-0107-0X
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 5948 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-09-017-947-1

Alignment Scores:

Pred. No.:	5.24e-84	Length:	5948
Score:	894.00	Matches:	361
Percent Similarity:	40.90%	Conservative:	295
Best Local Similarity:	22.51%	Mismatches:	672
Query Match:	12.17%	Indels:	276
DB:	4	Gaps:	56

US-10-020-095-4 (1-1428) x US-09-017-947-1 (1-5948)

```
Qy        18 AlaAlaLeuAlaVal---------------AlaProGlyProArgPheLeuValThrAla 32
             |||||||||   :::                :::   |||        :::   :::
Db        31 GCTGCTCTATTGATTGGTTTTCCAGGGTCTTCTCATGGGGCTCTCTACACCCTCATCACC 90

Qy        33 ProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeu---GluHis 51
             |||   ::::::|||                  |||   ||||||            :::
Db        91 CCTGCTGTTTTGCGAACAGACACAGAAGAGCAAATTTTGGTGGAGGCCCATGGAGACAGT 150

Qy        52 CysProSerGlnValThrValLysAla------------GluLeuLeuLysThrAlaSer 67
                |||   |||:::   :::                        ||||||
Db       151 ACTCCAAAACAGCTTGACATCTTTGTTCATGATTTTCCACGGAAGCAGAAAACCTTGTTC 210

Qy        68 AsnLeuThrValSerValLeuGluAlaGluGlyValPheGluLysGlySerPheLysThr 87
                         |||   :::      |||   |||:::                  |||
Db       211 CAAACCAGAGTAGATATGAATCCAGCAGGAGGCATGCTTGTCACT---------CCAACT 261

Qy        88 LeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyrGluLeuArgValThr 107
             :::   :::|||:::         :::::::::                 :::      |||
Db       262 ATAGAGATTCCAGCAAAAGAAGTGAGTACGGACTCCAGGCAAAATCAATATGTGGTTGTG 321

Qy       108 GlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThrLys 127
                   |||      ::::::                        ||||||:::::::::
Db       322 CAAGTAACTGGTCCTCAAGTGAGATTGGAAAAGGTGGTTCTCCTTTCTTACCAGAGTAGC 381
```

```
Qy    128 ArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluValLys 147
Db    382 ------TTTCTGTTTATCCAGACAGATAAAGGCATCTATACACCAGGGTCTCCAGTACTC 435

Qy    148 PheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIleLeu 167
Db    436 TATCGTGTTTTTCTATGGATCACAACACAAGCAAGATGAACAAAACTGTGATTGTTGAG 495

Qy    168 IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAspLeuGly 187
Db    496 TTTCAGACTCCAGAAGGCATTCTTGTCAGT---------TCTAATTCAGTTGACCTAAAC 546

Qy    188 ValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSerIle--- 206
Db    547 TTCTTC---TGGCCTTACAATTTACCAGACCTTGTCAGTTTGGGGACTTGGAGGATTGTG 603

Qy    207 ---GlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrVal 225
Db    604 GCCAAATATGAACATTCCCCAGAGAATTATACTGCATATTTTGATGTCAGGAAATATGTG 663

Qy    226 LeuProLysPheGluValThrLeuGlnThrPro---------LeuTyrCysSerMetAsn 242
Db    664 TTGCCAAGCTTTGAAGTCCGTCTGCAACCATCAGAGAAGTTTTTTACATTGACGGCAAT 723

Qy    243 SerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGly 262
Db    724 ---GAAAATTTCCACGTGTCTATCACTGCAAGGTACTTGTATGGAGAGGAAGTGGAAGGT 780

Qy    263 AspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle------ 280
Db    781 ---GTGGCCTTTGTCCTCTTTGGAGTGAAAATAGATGATGCTAAAAAGAGTATTCCAGAC 837

Qy    281 ------ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGlu 298
Db    838 TCACTCACGAGAATTCCGATTATTGATGGAGATGGGAAAGCAACACTAAAAAGAGATACA 897

Qy    299 MetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerPro 318
Db    898 TTCCGTTCTCGATTTCCAAATCTCAATGAGCTTGTTGGGCATACTCTGTATGCATCT--- 954

Qy    319 GlyProValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnVal 338
Db    955 ---------------GTAACAGTCATGACAGAATCAGGCAGTGATATGGTAGTGACTGAG 999

Qy    339 SerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThr 358
Db   1000 CAAAGCGGCATTCATATTGTGGCATCTCCCTATCAGATCCACTTCACAAAAACCCCCAAA 1059

Qy    359 ValLeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn 378
Db   1060 TATTTCAAGCCAGGAATGCCATATGAACTGACGGTGTATGTTACCAACCCTGATGGCTCA 1119

Qy    379 GlnLeuThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyr 398
Db   1120 CCAGCT------------------------GCCCATGTGCCAGTGGTATCAGAGGCCTTT 1155

Qy    399 ThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIle 418
Db   1156 ------------------CATTCTATGGGAACCACTTTGAGTGATGGGACTGCTAAGCTC 1197

Qy    419 AsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSer 438
Db   1198 ATCCTGAACATACCATTGAATGCTCAA---------AGCCTACCAATCACTGTTAGAACT 1248

Qy    439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db   1249 AACCATGGAGACCTCCCAAGAGAACGCCAGGCAACAAAGTCCATGACAGCCATAGCCTAC 1308

Qy    459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db   1309 CAAACCCAGGGAGGATCTGGAAACTATCTTCATGTAGCCATTACATCTACAGAGATTAAG 1368

Qy    479 ValGlySer------ProPheGluLeuValValSerGlyAsn------------LysArg 492
Db   1369 CCCGGAGATAACTTACCTGTCAATTTCAATGTGAAGGGCAATGCAAATTCACTGAAGCAG 1428

Qy    493 LeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLys--- 511
Db   1429 ATCAAATATTTCACATACCTCATATTGAATAAAGGGAAGATTTTCAAGGTTGGCAGGCAA 1488

Qy    512 ---------------GlnAsnSerThrMetPheSerLeuThrProGluAsnSerTrpThr 526
Db   1489 CCCAGGAGAGATGGGCAGAATCTGGTGACCATGAATCTGCATATCACTCCAGATCTCATC 1548

Qy    527 ProLysAlaCysValIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspVal 546
Db   1549 CCTTCCTTCCGGTTTGTGGCTTACTACCAAGTGGGAAACAACGAAATTGTGGCTGATTCT 1608

Qy    547 LeuLysIleProValGln------------------LeuValPheLys------AsnLys 558
Db   1609 GTCTGGGTGGATGTGAAGGATACCTGCATGGGAACGTTGGTTGTGAAAGGAGACAATCTA 1668

Qy    559 IleLysLeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIle 578
Db   1669 ATACAAATG------------------------CCAGGAGCTGCAATGAAAATCAAATTG 1704

Qy    579 SerValThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeu 598
Db   1705 GAA---GGGGATCCAGGTGCTCGGGTTGGTCTTGTGGCTGTGGACAAAGCAGTATATGTT 1761

Qy    599 MetAsnAlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsn 618
Db   1762 CTCAATGATAAATATAAGATTAGCCAAGCTAAGATATGGGACACAATAGAAAAGAGTGAC 1821

Qy    619 ThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeu 638
Db   1822 TTTGGCTGTACAGCTGGCAGTGGCCAGAATAATCTGGGTGTGTTTGAAGATGCTGGACTG 1881

Qy    639 TrpValLeuThrAspAlaAsnLeu---ThrLysAsp------------------------ 649
Db   1882 GCTCTGACAACCAGCACTAATCTCAACACCAAACAGAGATCAGCTGCAAAGTGTCCTCAG 1941

Qy    649 ------------------------------------------------------------ 649
Db   1942 CCTGCAAATCGGAGGCGTCGCAGTTCTGTTTTGCTGCTTGACAGCAACGCAAGCAAAGCG 2001

Qy    649 ------------------------------------------------------------ 649
Db   2002 GCAGAATTTCAGGATCAAGACCTGCGTAAATGCTGTGAAGATGTCATGCATGAGAACCCC 2061

Qy    649 ------------------------------------------------------------ 649
Db   2062 ATGGGGTACACTTGTGAAAAGCGTGCAAAATACATCCAGGAGGGAGATGCTTGTAAGGCT 2121

Qy    650 --------------------TyrIleAspGlyValTyrAsp---AsnAlaGluTyrAla 661
Db   2122 GCCTTCCTTGAATGCTGTCGCTACATCAAGGGGGTCCGAGATGAAAACCAACGGGAGAGC 2181

Qy    662 GluArgPheMet------GluGluAsnGlu---GlyHisIleValAspIleHisAspPhe 678
Db   2182 GAGTTGTTTCTGGCAAGAGATGATAATGAAGATGGTTTCATAGCAGATAGTGATATCATC 2241

Qy    679 SerLeuGlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeu--- 697
Db   2242 TCA---------------------AGGTCTGATTTCCCCAAGAGTTGGTTGTGGCTAACA 2280

Qy    698 ---------------AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThr 712
Db   2281 AAGGACTTGACCGAGGAGCCTAACAGTCAAGGGATTTCAAGCAAGACAATGTCTTTTTAT 2340

Qy    713 ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732
Db   2341 CTGAGGGATTCCATCACAACCTGGGTGGTGCTGGCTGTAAGCTTTACACCCACCAAAGGG 2400

Qy    733 LeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
```

```
Db   2401 ATC---TGTGTGGCTGAACCTTATGAAATAAGAGTCATGAAAGTCTTCTTCATTGATCTT 2457

Qy    753 AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772
Db   2458 CAAATGCCATATTCAGTAGTGAAGAATGAGCAGGTGGAGATTCGAGCTATTCTGCACAAC 2517

Qy    773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
Db   2518 TACGTTAACGAGGATATTTATGTGCGAGTGGAACTGTTATACAACCCAGCCTTC------ 2571

Qy    793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db   2572 ------TGCAGTGCTTCCACAAAAGGACAAAGATACCGACAGCAGTTCCCAATTAAAGCC 2625

Qy    813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly------GluIle 830
Db   2626 CTGTCCTCCAGAGCAGTACCGTTTGTGATAGTCCCATTAGAGCAAGGATTGCATGATGTT 2685

Qy    831 ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeu 850
Db   2686 GAGATTAAAGCAAGTGTCCAGGAAGCGTTGTGGTCAGACGGTGTGAGGAAGAAACTGAAA 2745

Qy    851 ValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870
Db   2746 GTTGTACCTGAAGGGGTACAGAAATCCATTGTGACTATTGTTAAACTGGACCCAAGGGCA 2805

Qy    871 AsnArgLeuGlnSerThr------------------LeuLysThrLeuSerPheSerPhe 884
Db   2806 AAAGGAGTTGGTGGAACACAGCTAGAAGTGATCAAAGCCCGCAAATTAGATGACAGAGTG 2865

Qy    885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904
Db   2866 CCTGACACAGAAATTGAAACCAAGATTATCATCCAAGGTGACCCTGTGGCTCAGATTATT 2925

Qy    905 GlyProSerIleAsnGly------LeuAlaSerLeuIleArgMetProTyrGlyCysGly 922
Db   2926 GAAAACTCAATTGATGGAAGTAAACTCAACCATCTCATTATCACTCCTTCTGGCTGTGGG 2985

Qy    923 GluGlnAsnMetIleAsnPheAlaProAsnIle---------TyrIleLeuAspTyrLeu 939
Db   2986 GAGCAAAATATGATCCGCATGGCCGCACCAGTTATTGCCACCTACTACCTGGACACCACA 3045

Qy    940 ThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGln 959
Db   3046 GAGCAGTGGGAGACTCTCGGCATAAATCGCAGGACTGAAGCTGTCAATCAGATCGTGACT 3105

Qy    960 GlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979
Db   3106 GGTTATGCCCAGCAGATGGTGTACAAGAAAGCAGATCATTCCTATGCAGCATTTACAAAC 3165

Qy    980 TyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPhe---LeuGlu 998
Db   3166 ------CGTGCATCTAGTTCTTGGCTAACAGCATATGTCGTAAAAGTCTTTGCCATGGCT 3219

Qy    999 AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeu--- 1017
Db   3220 GCCAAAATGGTAGCAGGCATTAGTCATGAAATCATTTGTGGAGGTGTGAGGTGGCTGATT 3279

Qy   1018 LysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGlu 1037
Db   3280 CTGAACAGGCAACAACCAGATGGAGCGTTCAAAGAAAATGCCCCTGTACTTTCTGGAACA 3339

Qy   1038 LeuGlnGly---------GlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db   3340 ATGCAGGGAGGAATTCAAGGTGCTGAAGAAGAAGTATATTTAACAGCTTTCATTCTGGTT 3399

Qy   1055 SerLeuLeuGlyTyrArg------------LysTyrGlnProAsnIleAsp------Val 1068
Db   3400 GCGTTGTTGGAATCCAAAACAATCTGCAATGACTATGTCAATAGTCTAGACAGCAGCATC 3459

Qy   1069 GlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db   3460 AAGAAGGCCACAAATTATTTACTCAAAAAGTATGAGAAA---CTGCAAAGGCCTTACACT 3516

Qy   1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeu 1108
Db   3517 ACAGCCCTCACAGCCTATGCTTTGGCT------------------GCTGCAGACCAACTC 3558

Qy   1109 AsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGlu 1128
Db   3559 AATGATGACAGGGTACTCATGGCAGCATCAACAGGAAGGGATCATTGG------------ 3606

Qy   1129 SerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAla 1148
Db   3607 ------------GAAGAATACAATGCTCACACCCACAACATTGAAGGCACTTCCTATGCC 3654

Qy   1149 LeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeu 1168
Db   3655 TTGTTGGCCCTGCTGAAAATGAAGAAATTTGATCAAACTGGTCCCATAGTCAGATGGCTG 3714

Qy   1169 SerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeu 1188
Db   3715 ACAGATCAGAATTTTTATGGGGAAACATATGGACAAACCCAAGCAACAGTTATGGCATTT 3774

Qy   1189 LysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg---ThrAsnIleGlnVal 1207
Db   3775 CAAGCTCTTGCTGAATATGAGATTCAGATGCCTACCCATAAGGACTTAAACTTAGATATT 3834

Qy   1208 ThrValThrGlyProSerSerProSerProLeu--------------------------- 1218
Db   3835 ACTATTGAACTGCCAGATCGAGAAGTACCTATAAGGTACAGAATTAATTATGAAAATGCT 3894

Qy   1219 ---------AlaValValGlnPro---------MetAlaValAsnIleSerAlaAsnGly 1232
Db   3895 CTCCTGGCTCGGACAGTAGAGACCAAACTCAACCAAGACATCACTGTGACAGCATCAGGT 3954

Qy   1233 PheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSer 1252
Db   3955 GATGGAAAAGCAACAATGACCATTTTGACATTCTATAACGCACAGTTG------------ 4002

Qy   1253 ArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGlu 1272
Db   4003 CAGGAGAAGGCAAATGTTTGCAATAAA------TTTCATCTTAATGTTTCTGTTGAA--- 4053

Qy   1273 AsnLysAspAspLeuAsnHisValAsp------------------LeuAsnValCysThr 1286
Db   4054 ---AACATCCACTTGAATGCAATGGGAGCCAAGGGAGCCCTCATGCTCAAGATCTGCACA 4110

Qy   1287 SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly 1306
Db   4111 AGGTATCTGGGAGAAGTTGATTCTACAATGACAATAATTGATATTTCTATGCTGACTGGT 4170

Qy   1307 PheMetValProSerGluAlaIleSer---------------LeuSerGluThrValLys 1321
Db   4171 TTTCTCCCTGATGCTGAAGACCTTACAAGGCTTTCTAAAGGAGTGGACAGATACATCTCC 4230

Qy   1322 LysValGluTyrAspHisGlyLysLeuAsn---------------LeuTyrLeuAspSer 1336
Db   4231 AGATATGAAGTTGACAATAATATGGCTCAGAAAGTAGCTGTTATCATTTACTTAAACAAG 4290

Qy   1337 ValAsnGluThrGln---PheCysValAsnIleProAlaValArgAsnPheLysValSer 1355
Db   4291 GTCTCCCACTCTGAAGATGAATGCCTGCACTTTAAGATTCTCAAGCATTTTGAAGTTGGC 4350

Qy   1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaVal 1375
Db   4351 TTCATTCAGCCAGGATCAGTCAAGGTGTACAGCTACTACAATCTAGATGAAAAATGTACC 4410

Qy   1376 ArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGly 1395
Db   4411 AAGTTCTACCATCCAGATAAAGGAACAGGCCTTCTCAATAAGATATGTATTGGTAACGTT 4470

Qy   1396 CysArgProCysGluAspGlyAlaSerGlySerHisHis--------------------- 1408
Db   4471 TGCCGATGTGCAGGAGAAACCTGTTCCTCGCTCAACCATCAGGAAAGGATTGATGTTCCA 4530
```

```
Qy       1409 ---------------------HisSerSerValIlePheIlePheCysPheLysLeuLeu 1421
                                    ::::::|||   :::::::::     |||||||||
Db       4531 TTACAAATTGAAAAAGCCTGCGAGACGAATGTGGATTATGTCTACAAAACCAAGCTGCTT 4590

Qy       1422 TyrPheMetGlu 1425
                       |||
Db       4591 CGAATAGAAGAA 4602

RESULT 9
US-08-447-411-44
; Sequence 44, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5924 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  4..69
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  70..4929
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  4..4929
US-08-447-411-44

Alignment Scores:
Pred. No.:              4.74e-83      Length:         5924
Score:                  885.00        Matches:        359
Percent Similarity:     40.84%        Conservative:   296
Best Local Similarity:  22.38%        Mismatches:     673
Query Match:            12.04%        Indels:         276
DB:                     1             Gaps:           56

US-10-020-095-4 (1-1428) x US-08-447-411-44 (1-5924)

Qy         18 AlaAlaLeuAlaVal---------------AlaProGlyProArgPheLeuValThrAla 32
              ||||||||    :::                  :::   |||     :::   :::
Db         31 GCTGCTCTATTGATTGGTTTTCCAGGGTCTTCTCATGGGGCTCTCTACACCCTCATCACC 90

Qy         33 ProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeu---GluHis 51
              |||   ::::::|||                    |||   ||||||         :::
Db         91 CCTGCTGTTTTGCGAACAGACACAGAAGAGCAAATTTTGGTGGAGGCCCATGGAGACAGT 150

Qy         52 CysProSerGlnValThrValLysAla------------GluLeuLeuLysThrAlaSer 67
                 |||   |||:::   :::                           ||||||
Db        151 ACTCCAAAACAGCTTGACATCTTTGTTCATGATTTTCCACGGAAGCAGAAAACCTTGTTC 210

Qy         68 AsnLeuThrValSerValLeuGluAlaGluGlyValPheGluLysGlySerPheLysThr 87
                    |||   :::      |||   |||:::                      |||
Db        211 CAAACCAGAGTAGATATGAATCCAGCAGGAGGCATGCTTGTCACT---------CCAACT 261

Qy         88 LeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyrGluLeuArgValThr 107
              :::   :::|||:::      :::::::::                 :::      |||
Db        262 ATAGAGATTCCAGCAAAAGAAGTGAGTACGGACTCCAGGCAAAATCAATATGTGGTTGTG 321

Qy        108 GlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThrLys 127
                    |||         ::::::                     ||||||:::::::::
Db        322 CAAGTAACTGGTCCTCAAGTGAGATTGGAAAAGGTGGTTCTCCTTTCTTACCAGAGTAGC 381

Qy        128 ArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluValLys 147
                    :::||||||||||||||||||      :::|||   |||         |||
Db        382 ------TTTCTGTTTATCCAGACAGATAAAGGCATCTATACACCAGGGTCTCCAGTACTC 435

Qy        148 PheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIleLeu 167
              :::|||:::   ::::::         :::               ::::::   :::
Db        436 TATCGTGTTTTTCTATGGATCACAACACAAGCAAGATGAACAAAACTGTGATTGTTGAG 495

Qy        168 IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAspLeuGly 187
                 :::   |||:::      |||:::               |||         ||||||
Db        496 TTTCAGACTCCAGAAGGCATTCTTGTCAGT---------TCTAATTCAGTTGACCTAAAC 546

Qy        188 ValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSerIle--- 206
                       :::   |||                  ||||||   |||   |||
Db        547 TTCTTC---TGGCCTTACAATTTACCAGACCTTGTCAGTTTGGGGACTTGGAGGATTGTG 603

Qy        207 ---GlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrVal 225
                 :::   :::   :::   :::   |||         |||   |||   :::||||||
Db        604 GCCAAATATGAACATTCCCCAGAGAATTATACTGCATATTTTGATGTCAGGAAATATGTG 663

Qy        226 LeuProLysPheGluValThrLeuGlnThrPro---------LeuTyrCysSerMetAsn 242
              ||||||   |||||||||   ||||||                  |||         |||
Db        664 TTGCCAAGCTTTGAAGTCCGTCTGCAACCATCAGAGAAGTTTTTTACATTGACGGCAAT 723

Qy        243 SerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGly 262
                 ::::::   :::   :::|||||||||:::|||   ||||||:::   |||:::|||
Db        724 ---GAAAATTTCCACGTGTCTATCACTGCAAGGTACTTGTATGGAGAGGAAGTGGAAGGT 780

Qy        263 AspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle------ 280
                 |||                  :::               ||||||:::|||
Db        781 ---GTGGCCTTTGTCCTCTTTGGAGTGAAAATAGATGATGCTAAAAAGAGTATTCCAGAC 837

Qy        281 ------ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGlu 298
                    |||:::         |||:::|||               :::         :::
Db        838 TCACTCACGAGAATTCCGATTATTGATGGAGATGGGAAAGCAACACTAAAAAGAGATACA 897

Qy        299 MetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerPro 318
                 ::::::         :::   |||   |||         :::         :::|||
Db        898 TTCCGTTCTCGATTTCCAAATCTCAATGAGCTTGTTGGGCATACTCTGTATGCATCT--- 954

Qy        319 GlyProValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnVal 338
                       :::|||   :::|||||||||   :::   :::
Db        955 ---------------GTAACAGTCATGACAGAATCAGGCAGTGATATGGTAGTGACTGAG 999

Qy        339 SerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThr 358
                 :::   :::                  |||   |||   |||
Db       1000 CAAAGCGGCATTCATATTGTGGCATCTCCCTATCAGATCCACTTCACAAAAACCCCCAAA 1059
```

```
Qy     359 ValLeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn 378
                  ||||||   :::   :::      ||||||   ||||||      ||||||:::
Db    1060 TATTTCAAGCCAGGAATGCCATATGAACTGACGGTGTATGTTACCAACCCTGATGGCTCA 1119

Qy     379 GlnLeuThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyr 398
                                               :::   |||            :::
Db    1120 CCAGCT------------------------GCCCATGTGCCAGTGGTATCAGAGGCCTTT 1155

Qy     399 ThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIle 418
                                :::   |||            :::         |||:::
Db    1156 ------------------CATTCTATGGGAACCACTTTGAGTGATGGGACTGCTAAGCTC 1197

Qy     419 AsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSer 438
                    :::|||                           ||||||         :::
Db    1198 ATCCTGAACATACCATTGAATGCTCAA---------AGCCTACCAATCACTGTTAGAACT 1248

Qy     439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
           :::         |||                  :::||||||      |||:::
Db    1249 AACCATGGAGACCTCCCAAGAGAACGCCAGGCAACAAAGTCCATGACAGCCATAGCCTAC 1308

Qy     459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
                 :::      |||      |||:::   :::                  |||||
Db    1309 CAAACCCAGGGAGGATCTGGAAACTATCTTCATGTAGCCATTACATCTACAGAGATTAAG 1368

Qy     479 ValGlySer------ProPheGluLeuValValSerGlyAsn------------LysArg 492
              |||         |||   :::            |||               |||:::
Db    1369 CCCGGAGATAACTTACCTGTCAAATTTCAATGTGAAGGGCAATGCAATTCACTGAAGCAG 1428

Qy     493 LeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLys--- 511
           :::|||      :::|||::::::::::::::|||::::::      ||||||:::
Db    1429 ATCAAATATTTCACATACCTCATATTGAATAAAGGGAAGATTTTCAAGGTTGGCAGGCAA 1488

Qy     512 ---------------GlnAsnSerThrMetPheSerLeuThrProGluAsnSerTrpThr 526
                          ||||||            :::|||
Db    1489 CCCAGGAGAGATGGGCAGAATCTGGTGACCATGAATCTGCATATCACTCCAGATCTCATC 1548

Qy     527 ProLysAlaCysValIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspVal 546
           |||            :::   ||||||         :::   ||||||::::::|||
Db    1549 CCTTCCTTCCGGTTTGTGGCTTACTACCAAGTGGGAAACAACGAAATTGTGGCTGATTCT 1608

Qy     547 LeuLysIleProValGln------------------LeuValPheLys------AsnLys 558
           :::   :::   |||:::                  ||||||   |||      |||
Db    1609 GTCTGGGTGGATGTGAAGGATACCTGCATGGGAACGTTGGTTGTGAAAGGAGACAATCTA 1668

Qy     559 IleLysLeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIle 578
           |||::::::                        |||         :::   :::::::::
Db    1669 ATACAAATG------------------------CCAGGAGCTGCAATGAAAATCAAATTG 1704

Qy     579 SerValThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeu 598
                       |||   :::   ||||||:::||||||||||||||||:::|||   :::
Db    1705 GAA---GGGGATCCAGGTGCTCGGGTTGGTCTTGTGGCTGTGGACAAAGCAGTATATGTT 1761

Qy     599 MetAsnAlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsn 618
           :::|||            |||:::         :::         :::|||      :::
Db    1762 CTCAATGATAAATATAAGATTAGCCAAGCTAAGATATGGGACACAATAGAAAAGAGTGAC 1821

Qy     619 ThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeu 638
              |||         |||         |||:::      ||||||::::::   ||||||
Db    1822 TTTGGCTGTACAGCTGGCAGTGGCCAGAATAATCTGGGTGTGTTTGAAGATGCTGGACTG 1881

Qy     639 TrpValLeuThrAspAlaAsnLeu---ThrLysAsp------------------------ 649
              :::   |||      ||||||   ||||||
Db    1882 GCTCTGACAACCAGCACTAATCTCAACACCAAACAGAGATCAGCTGCAAAGTGTCCTCAG 1941

Qy     649 ------------------------------------------------------------ 649

Db    1942 CCTGCAAATCGGAGGCGTCGCAGTTCTGTTTTGCTGCTTGACAGCAACGCAAGCAAAGCG 2001

Qy     649 ------------------------------------------------------------ 649

Db    2002 GCAGAATTTCAGGATCAAGACCTGCGTAAATGCTGTGAAGATGTCATGCATGAGAACCCC 2061

Qy     649 ------------------------------------------------------------ 649

Db    2062 ATGGGGTACACTTGTGAAAAGCGTGCAAAATACATCCAGGAGGGAGATGCTTGTAAGGCT 2121

Qy     650 ---------------------TyrIleAspGlyValTyrAsp---AsnAlaGluTyrAla 661
                                ||||||  ||||||    |||    |||         :::
Db    2122 GCCTTCCTTGAATGCTGTCGCTACATCAAGGGGGTCCGAGATGAAAACCAACGGGAGAGC 2181

Qy     662 GluArgPheMet------GluGluAsnGlu---GlyHisIleValAspIleHisAspPhe 678
           |||  |||:::         ::::::||||||    |||   |||  |||
Db    2182 GAGTTGTTTCTGGCAAGAGATGATAATGAAGATGGTTTCATAGCAGATAGTGATATCATC 2241

Qy     679 SerLeuGlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeu--- 697
           |||                     |||      ||||||::::::|||:::||||||
Db    2242 TCA---------------------AGGTCTGATTTCCCCAAGAGTTGGTTGTGGCTAACA 2280

Qy     698 ---------------AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThr 712
                          :::   |||                  :::
Db    2281 AAGGACTTGACCGAGGAGCCTAACAGTCAAGGGATTTCAAGCAAGACAATGTCTTTTTAT 2340

Qy     713 ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732
           :::   ||||||||||||:::||||||                  :::         |||
Db    2341 CTGAGGGATTCCATCACAACCTGGGTGGTGCTGGCTGTAAGCTTTACACCCACCAAAGGG 2400

Qy     733 LeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
           :::               |||   |||::::::      :::   ||||||||    |||
Db    2401 ATC---TGTGTGGCTGAACCTTATGAAATAAGAGTCATGAAAGTCTTCTTCATTGATCTT 2457

Qy     753 AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772
              :::||||||||||||::::::   |||:::      :::         :::   |||
Db    2458 CAAATGCCATATTCAGTAGTGAAGAATGAGCAGGTGGAGATTCGAGCTATTCTGCACAAC 2517

Qy     773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
           |||:::   :::         |||:::|||   :::      :::      |||
Db    2518 TACGTTAACGAGGATATTTATGTGCGAGTGGAACTGTTATACAACCCAGCCTTC------ 2571

Qy     793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
                    |||:::            ::::::|||               :::   :::
Db    2572 ------TGCAGTGCTTCCACAAAAGGACAAAGATACCGACAGCAGTTCCCAATTAAAGCC 2625

Qy     813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly------GluIle 830
                             |||   |||   |||   |||      |||      :::::
Db    2626 CTGTCCTCCAGAGCAGTACCGTTTGTGATAGTCCCATTAGAGCAAGGATTGCATGATGTT 2685

Qy     831 ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeu 850
              |||      :::                  ||||||  |||   :::   :::
Db    2686 GAGATTAAAGCAAGTGTCCAGGAAGCGTTGTGGTCAGACGGTGTGAGGAAGAAACTGAAA 2745

Qy     851 ValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870
           |||      ||||||::::::||||||            :::   ||||||
Db    2746 GTTGTACCTGAAGGGGTACAGAAATCCATTGTGACTATTGTTAAACTGGACCCAAGGGCA 2805

Qy     871 AsnArgLeuGlnSerThr------------------LeuLysThrLeuSerPheSerPhe 884
                 :::      |||                     :::   |||
Db    2806 AAAGGAGTTGGTGGAACACAGCTAGAAGTGATCAAAGCCCGCAAATTAGATGACAGAGTG 2865

Qy     885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904
           |||         :::            :::|||            :::      ::::::
Db    2866 CCTGACACAGAAATTGAAACCAAGATTATCATCCAAGGTGACCCTGTGGCTCAGATTATT 2925

Qy     905 GlyProSerIleAsnGly------LeuAlaSerLeuIleArgMetProTyrGlyCysGly 922
                 ||||||:::|||      |||      ||||||      |||   |||||||||
Db    2926 GAAAACTCAATTGATGGAAGTAAACTCAACCATCTCATTATCACTCCTTCTGGCTGTGGG 2985

Qy     923 GluGlnAsnMetIleAsnPheAlaProAsnIle---------TyrIleLeuAspTyrLeu 939
           |||||||||||||||      |||   :::            |||   ||||||
Db    2986 GAGCAAAATATGATCCGCATGGCCGCACCAGTTATTGCCACCTACTACCTGGACACCACA 3045

Qy     940 ThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGln 959
              :::   :::   |||      |||   :::   :::|||::::::   :::
Db    3046 GAGCAGTGGGAGACTCTCGGCATAAATCGCAGGACTGAAGCTGTCAATCAGATCGTGACT 3105

Qy     960 GlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979
```

```
               ||||||   :::::::::::|||::::::   |||   |||::::::|||||   |||
Db        3106 GGTTATGCCCAGCAGATGGTGTACAAGAAAGCAGATCATTCCTATGCAGCATTTACAAAC 3165

Qy         980 TyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPhe---LeuGlu 998
                   :::   |||:::||||||:::|||:::|||::::::   |||     :::
Db        3166 ------CGTGCATCTAGTTCTTGGCTAACAGCATATGTCGTAAAAGTCTTTGCCATGGCT 3219

Qy         999 AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeu--- 1017
               |||              |||         ::::::              ||||||
Db        3220 GCCAAAATGGTAGCAGGCATTAGTCATGAAATCATTTGTGGAGGTGTGAGGTGGCTGATT 3279

Qy        1018 LysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGlu 1037
                           |||:::   :::|||   |||   :::         |||:::
Db        3280 CTGAACAGGCAACAACCAGATGGAGCGTTCAAAGAAAATGCCCCTGTACTTTCTGGAACA 3339

Qy        1038 LeuGlnGly---------GlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
               :::||||||         |||   :::      |||   |||||||||:::|||:::
Db        3340 ATGCAGGGAGGAATTCAAGGTGCTGAAGAAGAAGTATATTTAACAGCTTTCATTCTGGTT 3399

Qy        1055 SerLeuLeuGlyTyrArg------------LysTyrGlnProAsnIleAsp------Val 1068
               :::||||||   :::               |||      ::::::|||      :::
Db        3400 GCGTTGTTGGAATCCAAAACAATCTGCAATGACTATGTCAATAGTCTAGACAGCAGCATC 3459

Qy        1069 GlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThr 1088
               :::::::::::::|||   ::::::   :::   :::            ||||||
Db        3460 AAGAAGGCCACAAATTATTTACTCAAAAAGTATGAGAAA---CTGCAAAGGCCTTACACT 3516

Qy        1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeu 1108
                  ||||||   |||||||||:::                  |||   :::   |||
Db        3517 ACAGCCCTCACAGCCTATGCTTTGGCT------------------GCTGCAGACCAACTC 3558

Qy        1109 AsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGlu 1128
               |||                                 |||      |||
Db        3559 AATGATGACAGGGTACTCATGGCAGCATCAACAGGAAGGGATCATTGG------------ 3606

Qy        1129 SerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAla 1148
                              :::   :::         :::   :::||||||      :::||||||
Db        3607 ------------GAAGAATACAATGCTCACACCCACAACATTGAAGGCACTTCCTATGCC 3654

Qy        1149 LeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeu 1168
               ||||||:::                  :::      :::      ||||||:::|||||||||
Db        3655 TTGTTGGCCCTGCTGAAAATGAAGAAATTTGATCAAACTGGTCCCATAGTCAGATGGCTG 3714

Qy        1169 SerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeu 1188
               :::   |||                  :::      ||||||   |||   :::|||
Db        3715 ACAGATCAGAATTTTTATGGGGAAACATATGGACAAACCCAAGCAACAGTTATGGCATTT 3774

Qy        1189 LysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg---ThrAsnIleGlnVal 1207
               :::||||||:::|||:::               |||   |||   :::      |||:::   :::
Db        3775 CAAGCTCTTGCTGAATATGAGATTCAGATGCCTACCCATAAGGACTTAAACTTAGATATT 3834

Qy        1208 ThrValThrGlyProSerSerProSerProLeu--------------------------- 1218
               |||:::      |||            |||:::
Db        3835 ACTATTGAACTGCCAGATCGAGAAGTACCTATAAGGTACAGAATTAATTATGAAAATGCT 3894

Qy        1219 ---------AlaValValGlnPro---------MetAlaValAsnIleSerAlaAsnGly 1232
                           |||:::                  :::   ::::::|||:::|||
Db        3895 CTCCTGGCTCGGACAGTAGAGACCAAACTCAACCAAGACATCACTGTGACAGCATCAGGT 3954

Qy        1233 PheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSer 1252
                  |||   |||            :::      ||||||   :::
Db        3955 GATGGAAAAGCAACAATGACCATTTTGACATTCTATAACGCACAGTTG------------ 4002

Qy        1253 ArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGlu 1272
               :::   :::   ::::::   |||:::         |||   |||:::|||:::|||:::
Db        4003 CAGGAGAAGGCAAATGTTTGCAATAAA------TTTCATCTTAATGTTTCTGTTGAA--- 4053

Qy        1273 AsnLysAspAspLeuAsnHisValAsp------------------LeuAsnValCysThr 1286
                        ||||||   :::                        |||   :::||||||
Db        4054 ---AACATCCACTTGAATGCAATGGGAGCCAAGGGAGCCCTCATGCTCAAGATCTGCACA 4110

Qy        1287 SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly 1306
                  :::   |||         |||   |||   ::::::::::::::::|||:::|||
Db        4111 AGGTATCTGGGAGAAGTTGATTCTACAATGACAATAATTGATATTTCTATGCTGACTGGT 4170

Qy        1307 PheMetValProSerGluAlaIleSer---------------LeuSerGluThrValLys 1321
               |||:::      :::|||   ::::::                  :::         :::
Db        4171 TTTCTCCCTGATGCTGAAGACCTTACAAGGCTTTCTAAAGGAGTGGACAGATACATCTCC 4230

Qy        1322 LysValGluTyrAspHisGlyLysLeuAsn---------------LeuTyrLeuAspSer 1336
               :::   |||   |||:::                           :::||||||:::
Db        4231 AGATATGAAGTTGACAATAATATGGCTCAGAAAGTAGCTGTTATCATTTACTTAAACAAG 4290

Qy        1337 ValAsnGluThrGln---PheCysValAsnIleProAlaValArgAsnPheLysValSer 1355
               |||:::   ::::::      |||::::::         :::::::::|||:::|||
Db        4291 GTCTCCCACTCTGAAGATGAATGCCTGCACTTTAAGATTCTCAAGCATTTTGAAGTTGGC 4350

Qy        1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaVal 1375
                     |||      ||||||   :::      ||||||               :::
Db        4351 TTCATTCAGCCAGGATCAGTCAAGGTGTACAGCTACTACAATCTAGATGAAAAATGTACC 4410

Qy        1376 ArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGly 1395
               :::   |||:::   :::                     :::
Db        4411 AAGTTCTACCATCCAGATAAAGGAACAGGCCTTCTCAATAAGATATGTATTGGTAACGTT 4470

Qy        1396 CysArgProCysGluAspGlyAlaSerGlySerHisHis--------------------- 1408
               ||||||         :::         |||      :::|||
Db        4471 TGCCGATGTGCAGGAGAAACCTGTTCCTCGCTCAACCATCAGGAAAGGATTGATGTTCCA 4530

Qy        1409 ---------------------HisSerSerValIlePheIlePheCysPheLysLeuLeu 1421
                                    ::::::|||   :::::::::      |||||||||
Db        4531 TTACAAATTGAAAAAGCCTGCGAGACGAATGTGGATTATGTCTACAAAACCAAGCTGCTT 4590

Qy        1422 TyrPheMetGlu 1425
                        |||
Db        4591 CGAATAGAAGAA 4602
```

```
RESULT 10
US-08-447-411-75
; Sequence 75, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  75:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4138 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..4001
US-08-447-411-75

Alignment Scores:
Pred. No.:                6.68e-71       Length:          4138
Score:                    768.50         Matches:         306
Percent Similarity:       41.88%         Conservative:    212
Best Local Similarity:    24.74%         Mismatches:      465
Query Match:              10.46%         Indels:          255
DB:                       1              Gaps:            46

US-10-020-095-4 (1-1428) x US-08-447-411-75 (1-4138)

Qy        349 TyrIleIleGluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAla 368
              |||   |||   |||                       |||||    :::    :::
Db         66 TATCAGATCTACTTCACAAAAACCCCCAAATATTTCAAGCCAGGAATGCCATATGAACTG 125

Qy        369 ThrValLysValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsn 388
              ||||||   ||||||:::   ||||||:::
Db        126 ACGGTGTATGTTACCAAACCTGATGGCTCACCAGCTGCC--------------------- 164

Qy        389 ValValIleThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGly 408
                 |||      |||::::::      ::::::|||                     |||
Db        165 CATGTGCCAGTGGTATCAGAGGCCATTCATTCTGAG---------------------GGA 203

Qy        409 AsnGlnLysMetGluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThrPhe 428
                          :::      |||:::              ||||||:::
Db        204 ACCACTTTGAGTGATGGGACTGCTAAGCTCTTCCTGAACACACCACAAAATGCTCAA--- 260

Qy        429 LysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeu 448
                          ||||||            ::::::      |||
Db        261 ------AGCCTACCGATCACTGTTAGAACTAACCATGGAGACCTCCCAAGAGAACGCCAG 314

Qy        449 GlySerLysSerSerMetAlaValHisSerLeuPheLysSerProSerLysThrTyrIle 468
                    ||||||      |||                  :::      |||      |||:::
Db        315 GCAATAAAGTCCATGACAGCCACAGCCTACCAAACCCAGGGAGGATCTGGAAACTATCTT 374

Qy        469 GlnLeuLysThrArgAspGluAsnIleLysValGlySer------ProPheGluLeuVal 486
                 :::                  ||||||   |||            |||
Db        375 CATGTAGCCATTACATCTACAGAGATTAAGCCCGGAGATAACTTACCTGTCAATTTCAAT 434

Qy        487 ValSerGlyAsn------------LysArgLeuLysGluLeuSerTyrMetValValSer 502
              |||   ||||||                        ::::::|||:::::::::::::::
Db        435 GTGAGGGGCAATGCAAATTCACTGAACCAGATCAAATATTTCACATACCTCATACTGAAT 494

Qy        503 ArgGlyGlnLeuValAlaValGlyLysGlnAsnSer---------------------Thr 515
              :::|||::::::         ||||||:::|||:::                     |||
Db        495 AAAGGGAAGATTTTCAAGGTTGGCAGGCAACACAGGGGAGATGGGGAGAATCTGGTGACC 554

Qy        516 Met---PheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleValTyr 534
              |||         :::||||||:::                |||         :::   |||
Db        555 ATGAATCTACATATCACTCCAGAT------CTCATTCCTTCCTTCCGGTTTGTGGCTTAC 608

Qy        535 TyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGln------ 552
              |||         :::   ||||||::::::|||   :::   :::   |||:::
Db        609 TACCAAGTGGGAAACAATGAAATTGTGGCTGATTCTGTCTGGGTGGATGTGAAGGATACC 668

Qy        553 ------------LeuValPheLys------------------AsnLysIleLysLeuTyr 562
                          ||||||   |||                  |||:::|||::::::
Db        669 TGCATGGGAACGTTGGTTGTGAAAGGAGCGACTTCCAGAGACAATCGAATACAAATG--- 725

Qy        563 TrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerValThrGln 582
                                    |||         :::   :::::::::
Db        726 ---------------------CCAGGAGCTGCAATGAAAATCAAATTGGAA---GGGGAT 761

Qy        583 ProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSer 602
              |||   :::   :::|||:::||||||||||||||||:::         ::::::|||
Db        762 CCAGGTGCTTGGATTGGTCTTGTGGCTGTGGACAAAGCAGAATATGTTCTCAATGATAAA 821

Qy        603 AsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyr 622
                 |||:::         :::         :::|||         :::   |||
Db        822 TATAAGATTAGCCAAGCTAAGATATGGGACACAATAGAAAAGAGTGACTTTGGCTGTACA 881

Qy        623 LeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeuThr 642
                 |||         |||:::      ||||||::::::   ||||||      :::   |||
Db        882 GCTGGCAGTGGCCAGAATAATCTGGGTGTGTTTGAAGATGCTGGACTGGCTCTGACAACC 941

Qy        643 AspAlaAsnLeuThr--------------------------------------------- 647
                 ||||||
Db        942 AGCACTAATCTCAACACCAAACAGAGATCAGCTGCAAAGTGTCCTCAGCCTGCAAATCGG 1001

Qy        647 ------------------------------------------------------------ 647

Db       1002 AGGCGTCGCAGTTCTGTTTTGCTGCTTGACAGCAACGCAAGCAAAGCGGCACAGTTTCAG 1061

Qy        648 ---------------LysAspTyrIleAspGlyValTyrAspAsn--------------- 657
                             |||   ||||||:::::::::|||
Db       1062 GATCAAGACCTGCGTAAATGCTGTGAAGATGGCATGCATGAGAACCCCATGGGGCACACT 1121

Qy        658 AlaGluTyrAlaGluArgPheMetGluGluAsnGluGly--------------------- 670
                 |||      |||::::::::::::|||      :::
Db       1122 TGTGAAAAGCGTGAAAAATACATCCAGGAGGGAGATGCTTGTAAGGCTGCCTTCCTCGAA 1181

Qy        671 ---------HisIleValAspIleHisAsp------------------------------ 677
                       :::|||   |||   |||
Db       1182 TGCTGTCACTACATCAAAGGGATCCAAGATGACAATAAACGGGAGAGCGAGTTGTTTCTG 1241

Qy        678 ---------------------------PheSerLeuGlySerSerProHisValArgLys 688
                                         |||      |||:::            |||
Db       1242 GCAAGAAGTGATTTTGAAGATGATTTATTTGGAGAAGGTAACATCACCTCA---AGGTCT 1298

Qy        689 HisPheProGluThrTrpIleTrpLeu--------------------------AspThr 699
                 |||||||||:::|||:::||||||
Db       1299 GATTTTCCTGAGAGTTGGTTGTGGCTAATGGAGCAGCTGTCTGAACATCCTAACAGTAAA 1358

Qy        700 AsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSer 719
                 :::   :::|||                  :::      ||||||||||||:::
Db       1359 GGGATTTCAAGCAAGATAGTACCTTTTTAT---------CTGAGGGATTCCATCACAACC 1409

Qy        720 TrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThr--- 738
              |||                     |||   |||   :::||||||:::   |||
Db       1410 TGG-----------------------GAGTTGCTGGCTGTGGGCCTTTCACCCACCAAA 1445

Qy        739 ------------------ProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
                                |||   |||:::         :::   |||||||||   |||
Db       1446 GGGATCTGTGTGGCTGAACCTTATGAAATAACAGTCATGAAAGACTTCTTCATTGATCTT 1505

Qy        753 AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772
                 ||||||||||||||||::::::   |||:::         :::         ::::::|||
Db       1506 CAACTGCCGTATTCAGTAGTGAAGAATGAGCAGGTGAAAATTCGAGCTGTTTTGTACAAC 1565

Qy        773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
              |||               |||:::|||   :::      |||         |||
Db       1566 TACGCTGACAAGGATATTTATGTACGAGTGGAACTGTTATACAGCCCAGCCTTCTGCAGT 1625

Qy        793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
                    |||   |||:::                  :::::::::   |||   :::   :::
Db       1626 GCTTCCACAGAAAGTCAA------------AGATACCGAGAGCAGTTGCCAATTAAAGCC 1673

Qy        813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly------GluIle 830
                          |||   |||   |||   |||         |||      ::::::
Db       1674 CTGTCCTCCAGGGCAGTATCGTTTGTGATAGTCCCATTAGAGCAAGGATTGCATGATGTT 1733

Qy        831 ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeu 850
                 :::|||   :::                     ||||||   |||   :::   :::
Db       1734 GAGGTTACAGCAAGTGTCCAGGGAGAGTTGATGTCAGATGGTGTGAAGAAGAAACTGAAA 1793

Qy        851 ValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870
```

```
Db   1794 GTTGTACCTGAAGGGGAATGGAAAAGTATTGTTACTATTATTGAACTGGACCCACATACA 1853
Qy    871 AsnArgLeuGlnSerThr------------------LeuLysThrLeuSerPheSerPhe 884
Db   1854 AAAGGAATTGGTGGAACACAGGTAGAATTGGTCAAAGCCAATAAATTAAATGACAGGGTT 1913
Qy    885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904
Db   1914 CCTGATACGGAAATAGAAACCAAGATTACTATTCAAGGTGATCCTGTGGCTCAGACTATT 1973
Qy    905 GlyProSerIleAsnGly------LeuAlaSerLeuIleArgMetProTyrGlyCysGly 922
Db   1974 GAAAACTCAATTGATGGAAGTAAACTCAACCATCTCATTATCACTCCTTTTGGCTGTGGG 2033
Qy    923 GluGlnAsnMetIleAsnPhe---AlaProAsnIle------TyrIleLeuAspTyrLeu 939
Db   2034 GAGCAAAATATGATCCGCATGACTGCACCAGTTATTGCCACCTACTACCTGGACACCACA 2093
Qy    940 ThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGln 959
Db   2094 CAGCAGTGGGAGACTCTCGGCATAAATCGCAGGACTGAAGCTGTCAATCAGATCATGACT 2153
Qy    960 GlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979
Db   2154 GGTTATGCCCAGCAGTTGGTGTACAAGAAAGCAGACCATTCCTATGCAGCATTTACAAAC 2213
Qy    980 TyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPhe---LeuGlu 998
Db   2214 ------AGTGCATCTAGTTCTTGGCTAACAGCATATGTTGTAAAAATCTTTGCCTTGGCT 2267
Qy    999 AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeu--- 1017
Db   2268 GCCAAAATTGTAAAAGACATTAACCATGAAATCGTTTGTGGAGGTATGAGGTGGCTGATT 2327
Qy   1018 LysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGlu 1037
Db   2328 CTGAACAGGCAACGAACAGATGGAGTGTTCAGAGAAAACGCCCCTGTACTTTTTGGAACA 2387
Qy   1038 LeuGlnGlyGlyAsnLysSerProVal---------ThrLeuThrAlaTyrIleValThr 1054
Db   2388 ATGCAGGGAGGCATTCAAGGTGCTGAACCAGAAGGATCTTTAACAGCTTTCATTCTGGTT 2447
Qy   1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
Db   2448 GCGTTGTTGGAATCCAGATCAATC---------------TGCAATGCATATATCAATATT 2492
Qy   1075 LeuGluSerGluPheSerArgGly------------------------------IleSer 1084
Db   2493 CTAGACAGCAGCATCAGTAAGGCCACAGATTATTTACTCAAAAAGTATGAGAAACTGCAA 2552
Qy   1085 AspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAla 1104
Db   2553 AGGCCTTACACTACAGCCCTCACAGCCTATGCTTTGGCT------------------GCT 2594
Qy   1105 LysGluAlaLeuAsn-Met------LeuThrTrpArgAlaGluGlnGluGlyGlyMetGl 1122
Db   2595 GCAGAACGACTCAATGATGACAGGGTACTCATGGCAGCATCAACAGGAAGGA-------- 2646
Qy   1122 nPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAspIl 1142
Db   2647 ---------------------ATCGTTGGGAAGAACCTAACGCCCACACCCATAAC-AT 2683
Qy   1142 eGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIl 1162
Db   2684 TGAAGGCACTTCCTATGCCTTGTTGGCCCTGCTGAAAATGAAGAAATTTGTTGAGGCCGG 2743
Qy   1162 eProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGl 1182
Db   2744 TCCTGTAGTCCAATGGCTGATAGATCAGCAATATTATGGGGGAACATATGGACAAACCCA 2803
Qy   1182 nAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluAr 1202
Db   2804 AGCAACAGTTATGATGTTTCAAGCTCTTGCTGAATATGAGATTCAGATGCCTACCCATAA 2863
Qy   1202 g---ThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----------- 1217
Db   2864 GGACTTAAACTTAGATATTACTATTGAACTGCCAGATCGAGAAGTACCTATAAGGTACAG 2923
Qy   1218 ----------------------LeuAlaValValGlnProMetAlaValAsn-------- 1227
Db   2924 AATTAATTATGAAAATGCTCTCCTGGCTCAGACAGTAGAGACCAAACTCAACGAAGACTT 2983
Qy   1228 ----IleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnVa 1246
Db   2984 CACTGTGTCAGCATCAGGTGATGGAAAAGCAACAATGACCATTTGACGGTCTATAATGC 3043
Qy   1246 lLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLe 1266
Db   3044 ACAATTG------------AGGGAGGATGCAAATGTTTGCAACAAA------TTCCATCT 3085
Qy   1266 uAspValAlaVal--------------------LysGluAsnLysAspAspLeuAsnHi 1279
Db   3086 TGATGTTTCTGTTGAAAACGTCCAGTTGAACTTAAAAGAGGCAAAGGGAGCCAAGGGAGC 3145
Qy   1279 sValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMe 1299
Db   3146 CCTCAAGCTCAAAATCTGCACTAGGTATCTGGGAGAAGTTGATTCTACAATGACAATAAT 3205
Qy   1299 tGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer---LeuSerGl 1318
Db   3206 TGATGTTTCTATGCTGACTGGTTTTGTCCCTGATACTGAAGACCTTACGAGGCTTTCTAA 3265
Qy   1318 uThrValLysLys------------ValGluTyrAspHis---------------GlyLy 1329
Db   3266 AGGAGTCGACAGATATATCTCCATGTTTGAAATTAACAATAATATGGCTCAGAAAGGAAC 3325
Qy   1329 sLeuAsnLeuTyrLeuAspSerValAsnGluThrGln---PheCysValAsnIleProAl 1348
Db   3326 TGTTATCATTTACTTAGACAAGGTCTCCCACTCTGAAGATGAATGCCTGCACTTTAAGAT 3385
Qy   1348 aValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTy 1368
Db   3386 TCTCAAGCATTTTGAAGTTGGCTTCATTCAGCCAGGATCAGTCAAGGTGTACAGCTACTA 3445
Qy   1368 rGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCysAs 1388
Db   3446 CAATCTAGATGAAAAATGTACCAAGATCTACCATCCAGATGAAGCAACAGGCCTTCTCAA 3505
Qy   1388 pLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAlaSer 1404
Db   3506 TAAGATATGTGTTGGTAACGTTTGCCGATGTGCAGAAGAAACCTGTTCC 3554
```

RESULT 11
US-08-662-227-33
; Sequence 33, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER:  US/08/662,227
;     FILING DATE:  14-JUN-1996
;     CLASSIFICATION:  530
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  OBLON, NORMAN F.
;     REGISTRATION NUMBER:  24,618
;     REFERENCE/DOCKET NUMBER:  1126-0107-0X
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  703-413-3000
;     TELEFAX:  703-413-2220
; INFORMATION FOR SEQ ID NO:  33:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  4138 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
US-08-662-227-33

Alignment Scores:
Pred. No.:              6.68e-71      Length:         4138
Score:                  768.50        Matches:        306
Percent Similarity:     41.88%        Conservative:   212
Best Local Similarity:  24.74%        Mismatches:     465
Query Match:            10.46%        Indels:         255
DB:                     2             Gaps:           46

US-10-020-095-4 (1-1428) x US-08-662-227-33 (1-4138)

Qy    349 TyrIleIleGluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAla 368
Db     66 TATCAGATCTACTTCACAAAAACCCCAAATATTTCAAGCCAGGAATGCCATATGAACTG 125

Qy    369 ThrValLysValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsn 388
Db    126 ACGGTGTATGTTACCAAACCTGATGGCTCACCAGCTGCC--------------------- 164

Qy    389 ValValIleThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGly 408
Db    165 CATGTGCCAGTGGTATCAGAGGCCATTCATTCTGAG---------------------GGA 203

Qy    409 AsnGlnLysMetGluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThrPhe 428
Db    204 ACCACTTTGAGTGATGGGACTGCTAAGCTCTTCCTGAACACACCACAAAATGCTCAA--- 260

Qy    429 LysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeu 448
Db    261 ------AGCCTACCGATCACTGTTAGAACTAACCATGGAGACCTCCCAAGAGAACGCCAG 314

Qy    449 GlySerLysSerSerMetAlaValHisSerLeuPheLysSerProSerLysThrTyrIle 468
Db    315 GCAATAAAGTCCATGACAGCCACAGCCTACCAAACCCAGGGAGGATCTGGAAACTATCTT 374

Qy    469 GlnLeuLysThrArgAspGluAsnIleLysValGlySer------ProPheGluLeuVal 486
Db    375 CATGTAGCCATTACATCTACAGAGATTAAGCCCGGAGATAACTTACCTGTCAATTTCAAT 434

Qy    487 ValSerGlyAsn------------LysArgLeuLysGluLeuSerTyrMetValValSer 502
Db    435 GTGAGGGGCAATGCAAATTCACTGAACCAGATCAAATATTTCACATACCTCATACTGAAT 494

Qy    503 ArgGlyGlnLeuValAlaValGlyLysGlnAsnSer---------------------Thr 515
Db    495 AAAGGGAAGATTTTCAAGGTTGGCAGGCAACACAGGGGAGATGGGGAGAATCTGGTGACC 554

Qy    516 Met---PheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleValTyr 534
Db    555 ATGAATCTACATATCACTCCAGAT------CTCATTCCTTCCTTCCGGTTTGTGGCTTAC 608

Qy    535 TyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGln------ 552
Db    609 TACCAAGTGGGAAACAATGAAATTGTGGCTGATTCTGTCTGGGTGGATGTGAAGGATACC 668

Qy    553 ------------LeuValPheLys------------------AsnLysIleLysLeuTyr 562
Db    669 TGCATGGGAACGTTGGTTGTGAAAGGAGCGACTTCCAGAGACAATCGAATACAAATG--- 725

Qy    563 TrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerValThrGln 582
Db    726 ---------------------CCAGGAGCTGCAATGAAAATCAAATTGGAA---GGGGAT 761

Qy    583 ProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSer 602
Db    762 CCAGGTGCTTGGATTGGTCTTGTGGCTGTGGACAAAGCAGAATATGTTCTCAATGATAAA 821

Qy    603 AsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyr 622
Db    822 TATAAGATTAGCCAAGCTAAGATATGGGACACAATAGAAAAGAGTGACTTTGGCTGTACA 881

Qy    623 LeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeuThr 642
Db    882 GCTGGCAGTGGCCAGAATAATCTGGGTGTGTTTGAAGATGCTGGACTGGCTCTGACAACC 941

Qy    643 AspAlaAsnLeuThr--------------------------------------------- 647
Db    942 AGCACTAATCTCAACACCAAACAGAGATCAGCTGCAAAGTGTCCTCAGCCTGCAAATCGG 1001

Qy    647 ------------------------------------------------------------ 647
Db   1002 AGGCGTCGCAGTTCTGTTTTGCTGCTTGACAGCAACGCAAGCAAAGCGGCACAGTTTCAG 1061

Qy    648 ---------------LysAspTyrIleAspGlyValTyrAspAsn--------------- 657
Db   1062 GATCAAGACCTGCGTAAATGCTGTGAAGATGGCATGCATGAGAACCCCATGGGGCACACT 1121

Qy    658 AlaGluTyrAlaGluArgPheMetGluGluAsnGluGly--------------------- 670
Db   1122 TGTGAAAAGCGTGAAAAATACATCCAGGAGGGAGATGCTTGTAAGGCTGCCTTCCTCGAA 1181

Qy    671 ---------HisIleValAspIleHisAsp------------------------------ 677
Db   1182 TGCTGTCACTACATCAAAGGGATCCAAGATGACAATAAACGGGAGAGCGAGTTGTTTCTG 1241

Qy    678 --------------------------PheSerLeuGlySerSerProHisValArgLys 688
Db   1242 GCAAGAAGTGATTTTGAAGATGATTTATTTGGAGAAGGTAACATCACCTCA---AGGTCT 1298

Qy    689 HisPheProGluThrTrpIleTrpLeu--------------------------AspThr 699
Db   1299 GATTTTCCTGAGAGTTGGTTGTGGCTAATGGAGCAGCTGTCTGAACATCCTAACAGTAAA 1358

Qy    700 AsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSer 719
Db   1359 GGGATTTCAAGCAAGATAGTACCTTTTAT---------CTGAGGGATTCCATCACAACC 1409

Qy    720 TrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThr--- 738
Db   1410 TGG----------------------GAGTTGCTGGCTGTGGGCCTTTCACCCACCAAA 1445

Qy    739 ------------------ProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
Db   1446 GGGATCTGTGTGGCTGAACCTTATGAAATAACAGTCATGAAAGACTTCTTCATTGATCTT 1505

Qy    753 AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772
Db   1506 CAACTGCCGTATTCAGTAGTGAAGAATGAGCAGGTGAAAATTCGAGCTGTTTTGTACAAC 1565

Qy    773 TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
Db   1566 TACGCTGACAAGGATATTTATGTACGAGTGGAACTGTTATACAGCCCAGCCTTCTGCAGT 1625

Qy    793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db   1626 GCTTCCACAGAAAGTCAA------------AGATACCGAGAGCAGTTGCCAATTAAAGCC 1673
```

```
Qy        813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly------GluIle 830
                        |||  |||  |||  |||        |||       ::::::
Db       1674 CTGTCCTCCAGGGCAGTATCGTTTGTGATAGTCCCATTAGAGCAAGGATTGCATGATGTT 1733

Qy        831 ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeu 850
                 :::|||   :::                  |||||  |||   :::   :::
Db       1734 GAGGTTACAGCAAGTGTCCAGGGAGAGTTGATGTCAGATGGTGTGAAGAAGAAACTGAAA 1793

Qy        851 ValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870
              |||      |||||       |||||             |||   |||||
Db       1794 GTTGTACCTGAAGGGGAATGGAAAAGTATTGTTACTATTATTGAACTGGACCCACATACA 1853

Qy        871 AsnArgLeuGlnSerThr------------------LeuLysThrLeuSerPheSerPhe 884
                 :::     |||                             |||:::
Db       1854 AAAGGAATTGGTGGAACACAGGTAGAATTGGTCAAAGCCAATAAATTAAATGACAGGGTT 1913

Qy        885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904
              |||         :::               :::|||      :::         :::
Db       1914 CCTGATACGGAAATAGAAACCAAGATTACTATTCAAGGTGATCCTGTGGCTCAGACTATT 1973

Qy        905 GlyProSerIleAsnGly------LeuAlaSerLeuIleArgMetProTyrGlyCysGly 922
                    ||||||:::|||      |||       ||||||    |||:::|||||||||
Db       1974 GAAAACTCAATTGATGGAAGTAAACTCAACCATCTCATTATCACTCCTTTTGGCTGTGGG 2033

Qy        923 GluGlnAsnMetIleAsnPhe---AlaProAsnIle------TyrIleLeuAspTyrLeu 939
              ||||||||||||||||||         ||||||  |||      |||   ||||||
Db       2034 GAGCAAAATATGATCCGCATGACTGCACCAGTTATTGCCACCTACTACCTGGACACCACA 2093

Qy        940 ThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGln 959
                 :::   :::   |||      |||   :::   :::|||::::::   :::
Db       2094 CAGCAGTGGGAGACTCTCGGCATAAATCGCAGGACTGAAGCTGTCAATCAGATCATGACT 2153

Qy        960 GlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn 979
              ||||||    ::::::|||:::|||::::::   |||    |||::::::||||||   |||
Db       2154 GGTTATGCCCAGCAGTTGGTGTACAAGAAAGCAGACCATTCCTATGCAGCATTTACAAAC 2213

Qy        980 TyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPhe---LeuGlu 998
                    :::   |||:::||||||:::|||:::|||::::::   |||    |||
Db       2214 ------AGTGCATCTAGTTCTTGGCTAACAGCATATGTTGTAAAAATCTTTGCCTTGGCT 2267

Qy        999 AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeu--- 1017
              |||          ||||||:::         ::::::             ||||||
Db       2268 GCCAAAATTGTAAAAGACATTAACCATGAAATCGTTTGTGGAGGTATGAGGTGGCTGATT 2327

Qy       1018 LysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGlu 1037
                       |||:::::::::|||   |||   :::         |||:::
Db       2328 CTGAACAGGCAACGAACAGATGGAGTGTTCAGAGAAAACGCCCCTGTACTTTTTGGAACA 2387

Qy       1038 LeuGlnGlyGlyAsnLysSerProVal---------ThrLeuThrAlaTyrIleValThr 1054
              :::||||||||||   :::                 :::||||||||||:::|||:::
Db       2388 ATGCAGGGAGGCATTCAAGGTGCTGAACCAGAAGGATCTTTAACAGCTTTCATTCTGGTT 2447

Qy       1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
              :::||||||   |||                             |||:::
Db       2448 GCGTTGTTGGAATCCAGATCAATC---------------TGCAATGCATATATCAATATT 2492

Qy       1075 LeuGluSerGluPheSerArgGly-----------------------------IleSer 1084
              |||:::|||   |||:::                                    :::
Db       2493 CTAGACAGCAGCATCAGTAAGGCCACAGATTATTTACTCAAAAAGTATGAGAAACTGCAA 2552

Qy       1085 AspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAla 1104
                    ||||||   ||||||   |||||||||:::                   |||
Db       2553 AGGCCTTACACTACAGCCCTCACAGCCTATGCTTTGGCT-----------------GCT 2594

Qy       1105 LysGluAlaLeuAsn-Met------LeuThrTrpArgAlaGluGlnGluGlyGlyMetGl 1122
                 |||   ||||||  |||          :::|||:::   :::|||||||||
Db       2595 GCAGAACGACTCAATGATGACAGGGTACTCATGGCAGCATCAACAGGAAGGA-------- 2646

Qy       1122 nPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAspIl 1142
                                   :::   :::      |||      :::   ||
Db       2647 ---------------------ATCGTTGGGAAGAACCTAACGCCCACACCCATAAC-AT 2683

Qy       1142 eGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIl 1162
               ||||      :::|||||||||||:::                :::      |||
Db       2684 TGAAGGCACTTCCTATGCCTTGTTGGCCCTGCTGAAAATGAAGAAATTTGTTGAGGCCGG 2743

Qy       1162 eProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGl 1182
                |||:::::::::||||||        |||:::          |||    :::      ||||
Db       2744 TCCTGTAGTCCAATGGCTGATAGATCAGCAATATTATGGGGAACATATGGACAAACCCA 2803

Qy       1182 nAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluAr 1202
               |   |||    :::      :::||||||:::|||:::            |||   |||    ::
Db       2804 AGCAACAGTTATGATGTTTCAAGCTCTTGCTGAATATGAGATTCAGATGCCTACCCATAA 2863

Qy       1202 g---ThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----------- 1217
               :      |||:::   :::|||:::          |||          |||
Db       2864 GGACTTAAACTTAGATATTACTATTGAACTGCCAGATCGAGAAGTACCTATAAGGTACAG 2923

Qy       1218 ---------------------LeuAlaValValGlnProMetAlaValAsn-------- 1227
                                   ||||||                   :::|||
Db       2924 AATTAATTATGAAAATGCTCTCCTGGCTCAGACAGTAGAGACCAAACTCAACGAAGACTT 2983

Qy       1228 ----IleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnVa 1246
                  :::||||||::::||  |||   |||         :::      ||||||||
Db       2984 CACTGTGTCAGCATCAGGTGATGGAAAAGCAACAATGACCATTTTGACGGTCTATAATGC 3043

Qy       1246 lLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLe 1266
                :::                |||          ::::::   |||:::      |||   ||
Db       3044 ACAATTG------------AGGGAGGATGCAAATGTTTGCAACAAA------TTCCATCT 3085

Qy       1266 uAspValAlaVal--------------------LysGluAsnLysAspAspLeuAsnHi 1279
               ||||||:::|||                       ||||||   |||
Db       3086 TGATGTTTCTGTTGAAAACGTCCAGTTGAACTTAAAAGAGGCAAAGGGAGCCAAGGGAGC 3145

Qy       1279 sValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMe 1299
                :::   |||   :::||||||    :::   |||          |||   |||   :::::
Db       3146 CCTCAAGCTCAAAATCTGCACTAGGTATCTGGGAGAAGTTGATTCTACAATGACAATAAT 3205

Qy       1299 tGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer---LeuSerGl 1318
                :::|||::::::|||:::||||||:::         :::|||    ::::::   ||||||::
Db       3206 TGATGTTTCTATGCTGACTGGTTTTGTCCCTGATACTGAAGACCTTACGAGGCTTTCTAA 3265

Qy       1318 uThrValLysLys------------ValGluTyrAspHis---------------GlyLy 1329
               :   |||   :::                |||   ::::::                    |||
Db       3266 AGGAGTCGACAGATATATCTCCATGTTTGAAATTAACAATAATATGGCTCAGAAAGGAAC 3325

Qy       1329 sLeuAsnLeuTyrLeuAspSerValAsnGluThrGln---PheCysValAsnIleProAl 1348
                :::   :::||||||||    |||:::   ::::::         |||::::::
Db       3326 TGTTATCATTTACTTAGACAAGGTCTCCCACTCTGAAGATGAATGCCTGCACTTTAAGAT 3385

Qy       1348 aValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTy 1368
                :::::::::|||:::|||            |||      ||||||   :::      |||||
Db       3386 TCTCAAGCATTTTGAAGTTGGCTTCATTCAGCCAGGATCAGTCAAGGTGTACAGCTACTA 3445

Qy       1368 rGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCysAs 1388
                             :::      :::   |||:::   :::                    ::
Db       3446 CAATCTAGATGAAAAATGTACCAAGATCTACCATCCAGATGAAGCAACAGGCCTTCTCAA 3505

Qy       1388 pLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAlaSer 1404
               :                     ||||||   |||:::      |||
Db       3506 TAAGATATGTGTTGGTAACGTTTGCCGATGTGCAGAAGAAACCTGTTCC 3554
```

RESULT 12
US-09-017-947-33
; Sequence 33, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
;   APPLICANT: VOGEL, CARL-WILHELM
;   APPLICANT: BREDEHORST, REINHORST
;   APPLICANT: KOCK, MICHAEL
;   APPLICANT: FRITZINGER, DAVID
;   TITLE OF INVENTION: RECOMBINANT PROCVF
;   NUMBER OF SEQUENCES: 39
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

```
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/017,947
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/662,227
;      FILING DATE:  14-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-0107-0X
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4138 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-017-947-33

Alignment Scores:
Pred. No.:              6.68e-71      Length:        4138
Score:                  768.50        Matches:       306
Percent Similarity:     41.88%        Conservative:  212
Best Local Similarity:  24.74%        Mismatches:    465
Query Match:            10.46%        Indels:        255
DB:                     4             Gaps:          46

US-10-020-095-4 (1-1428) x US-09-017-947-33 (1-4138)

Qy        349 TyrIleIleGluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAla 368
               |||   |||   |||                    ||||||    :::   :::
Db         66 TATCAGATCTACTTCACAAAAACCCCCAAATATTTCAAGCCAGGAATGCCATATGAACTG 125

Qy        369 ThrValLysValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsn 388
              ||||||   ||||||:::   ||||||:::
Db        126 ACGGTGTATGTTACCAAACCTGATGGCTCACCAGCTGCC--------------------- 164

Qy        389 ValValIleThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGly 408
                 |||   |||::::::   ::::::|||                         |||
Db        165 CATGTGCCAGTGGTATCAGAGGCCATTCATTCTGAG---------------------GGA 203

Qy        409 AsnGlnLysMetGluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThrPhe 428
                       :::         |||:::         ||||||:::
Db        204 ACCACTTTGAGTGATGGGACTGCTAAGCTCTTCCTGAACACACCACAAAATGCTCAA--- 260

Qy        429 LysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeu 448
                       ||||||      ::::::         |||
Db        261 ------AGCCTACCGATCACTGTTAGAACTAACCATGGAGACCTCCCAAGAGAACGCCAG 314

Qy        449 GlySerLysSerSerMetAlaValHisSerLeuPheLysSerProSerLysThrTyrIle 468
                    ||||||   |||            :::   |||      |||:::
Db        315 GCAATAAAGTCCATGACAGCCACAGCCTACCAAACCCAGGGAGGATCTGGAAACTATCTT 374

Qy        469 GlnLeuLysThrArgAspGluAsnIleLysValGlySer------ProPheGluLeuVal 486
              :::                     ||||||   |||            |||
Db        375 CATGTAGCCATTACATCTACAGAGATTAAGCCCGGAGATAACTTACCTGTCAATTTCAAT 434

Qy        487 ValSerGlyAsn------------LysArgLeuLysGluLeuSerTyrMetValValSer 502
              |||   ||||||                  ::::::|||          :::|||::::::::::::
Db        435 GTGAGGGGCAATGCAAATTCACTGAACCAGATCAAATATTTCACATACCTCATACTGAAT 494

Qy        503 ArgGlyGlnLeuValAlaValGlyLysGlnAsnSer--------------------Thr 515
              :::|||::::::   ||||||:::|||:::                          |||
Db        495 AAAGGGAAGATTTTCAAGGTTGGCAGGCAACACAGGGGAGATGGGGAGAATCTGGTGACC 554

Qy        516 Met---PheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleValTyr 534
              |||         :::||||||:::                |||            ::: |||
Db        555 ATGAATCTACATATCACTCCAGAT------CTCATTCCTTCCTTCCGGTTTGTGGCTTAC 608

Qy        535 TyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGln------ 552
              |||         :::   ||||||::::::|||   :::   :::   |||:::
Db        609 TACCAAGTGGGAAACAATGAAATTGTGGCTGATTCTGTCTGGGTGGATGTGAAGGATACC 668

Qy        553 ------------LeuValPheLys-----------------AsnLysIleLysLeuTyr 562
                          ||||||   |||                 |||:::|||::::::
Db        669 TGCATGGGAACGTTGGTTGTGAAAGGAGCGACTTCCAGAGACAATCGAATACAAATG--- 725

Qy        563 TrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerValThrGln 582
                                   |||         :::   :::::::::
Db        726 ---------------------CCAGGAGCTGCAATGAAAATCAAATTGGAA---GGGGAT 761

Qy        583 ProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSer 602
              |||   :::   :::|||:::||||||||||||||||||:::         ::::::|||
Db        762 CCAGGTGCTTGGATTGGTCTTGTGGCTGTGGACAAAGCAGAATATGTTCTCAATGATAAA 821

Qy        603 AsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyr 622
                    |||:::         :::         :::|||      :::   |||
Db        822 TATAAGATTAGCCAAGCTAAGATATGGGACACAATAGAAAAGAGTGACTTTGGCTGTACA 881

Qy        623 LeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeuThr 642
                 |||      |||:::         ||||||::::::   ||||||   :::   |||
Db        882 GCTGGCAGTGGCCAGAATAATCTGGGTGTGTTTGAAGATGCTGGACTGGCTCTGACAACC 941

Qy        643 AspAlaAsnLeuThr--------------------------------------------- 647
                    ||||||
Db        942 AGCACTAATCTCAACACCAAACAGAGATCAGCTGCAAAGTGTCCTCAGCCTGCAAATCGG 1001

Qy        647 ------------------------------------------------------------ 647

Db       1002 AGGCGTCGCAGTTCTGTTTTGCTGCTTGACAGCAACGCAAGCAAAGCGGCACAGTTTCAG 1061

Qy        648 ---------------LysAspTyrIleAspGlyValTyrAspAsn--------------- 657
                             |||         ||||||:::::::::|||
Db       1062 GATCAAGACCTGCGTAAATGCTGTGAAGATGGCATGCATGAGAACCCCATGGGGCACACT 1121

Qy        658 AlaGluTyrAlaGluArgPheMetGluGluAsnGluGly--------------------- 670
              |||      |||::::::::::::|||      :::
Db       1122 TGTGAAAAGCGTGAAAAATACATCCAGGAGGGAGATGCTTGTAAGGCTGCCTTCCTCGAA 1181

Qy        671 ---------HisIleValAspIleHisAsp------------------------------ 677
                        :::|||      |||   |||
Db       1182 TGCTGTCACTACATCAAAGGGATCCAAGATGACAATAAACGGGAGAGCGAGTTGTTTCTG 1241

Qy        678 ---------------------------PheSerLeuGlySerSerProHisValArgLys 688
                                         |||   |||:::            |||
Db       1242 GCAAGAAGTGATTTTGAAGATGATTTATTTGGAGAAGGTAACATCACCTCA---AGGTCT 1298

Qy        689 HisPheProGluThrTrpIleTrpLeu---------------------------AspThr 699
              ||||||||:::|||:::||||||
Db       1299 GATTTTCCTGAGAGTTGGTTGTGGCTAATGGAGCAGCTGTCTGAACATCCTAACAGTAAA 1358

Qy        700 AsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSer 719
                 :::   :::|||                  :::      ||||||||||||:::
Db       1359 GGGATTTCAAGCAAGATAGTACCTTTTTAT---------CTGAGGGATTCCATCACAACC 1409

Qy        720 TrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThr--- 738
              |||                        |||   |||   :::||||||:::   |||
Db       1410 TGG-----------------------GAGTTGCTGGCTGTGGGCCTTTCACCCACCAAA 1445

Qy        739 ------------------ProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
                                |||   |||:::         :::   |||||||||   |||
```

```
Qy        416 -------------GlnLysIleAsn------TyrThrValProGln------------- 424
                           :::::|||        ||||||
Db      16768 CGATGCCGACAATGAACGTATCGGTCGGCTTTATACCGCAGCTTCAGGTGTGGTGACGCC 16709

Qy        425 -----------------------------SerGlyThrPheLysIleGluPheProIleLe 435
                                           ||||||   :::::::::   |||
Db      16708 TGCTCCGGCCGAATGGGGCAAGTTGCCTTCGGGACAGTACCGATTGCGTTTTGGCGAGAG 16649

Qy        435 u------GluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSe 453
                     |||||||||         :::           |||           |||
Db      16648 TGGCGGAGAGGATTCCGACTTTGTTACGAAGGACGTTTATCTCTTCCGTCCGAAAGATCG 16589

Qy        453 rMetAlaValHisSerLeuPheLysSerProSerLys---------ThrTyrIleGlnLe 470
                                        :::||||||:::         ||||||:::
Db      16588 TCAG---------------TTGTCGAATCCCTCCCAAGGCCTTTGGACATATGTGGCGGA 16544

Qy        470 uLysThrArgAspGluAsn------------IleLysValGlySerPro----------- 482
               :::       |||    :::              |||    ||||||:::
Db      16543 GGAGAAATACGACAGTCACCGCCCTGCCCGCATCCTGGTCGGTGCTACTCCCAAGGATGC 16484

Qy        483 -------------------------PheGluLeuValValSerGlyAsnLysArgLeuLy 494
                                       ::::::|||       :::|||      |||
Db      16483 GGATGCTCAGGAGTTGTATCTGTTCTACGACCTGACTCAGGCAGGC------CGATTCAT 16430

Qy        494 sGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSe 514
                |||          :::          |||::::::|||    :::
Db      16429 CGAAAGGAAGATGATCGCTTGCCGGCCGGGTGAGATCGTCGAGATA------------GT 16382

Qy        514 rThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleValTy 534
                 |||          |||          |||::::::    :::
Db      16381 GCCGATGCTTCCGGCAGCTCCG-------------TTGCCGGAGAGTATGAATGTTTCTCT 16334

Qy        534 rTyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGlnLeuVa 554
                |||          |||   :::   |||      :::    :::       :::|||
Db      16333 CTATTATGTATATGGAGGGAGGCTGTATAGTCGAAGCATAGATCTGGAACGCAGACTGCC 16274

Qy        554 lPheLysAsnLysIleLysLeuTyrTrpSer------LysValLysAlaGluProSerGl 572
                 |||   :::||||||:::  ||||||        :::   |||       |||   ||
Db      16273 C---AAGCGAGAGATAAAGATGTCGTGGAGTACTTTCCGCGACAAGCTCCGACCTGGTGA 16217

Qy        572 uLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaVa 592
              ||||   :::    ::::::    :::       :::
Db      16216 GAAAGAGACATGGAAGCTCCGTCTCGCCGATGCGGAG-----------------GGCAA 16175

Qy        592 lAspLysSerValAsnLeuMetAsnAla---SerAsnAspIleThrMetGluAsnValVa 611
                    |||      :::|||  |||          |||      ::::::   :::||
Db      16174 ACCGCTCTCTCATACCATGATGGCTGCTTGGATGTACGATGCTGCTCTTGATAAGATCGT 16115

Qy        611 lHis--------------------------------------------------------- 612
              ||||
Db      16114 GCACAGTTCTGTTTTCTCCTCCCGATTCTTCTCTTTTGCCGACTCTCCGGCTTCTTTGGT 16055

Qy        613 -GluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSer-------- 629
                 |||          |||:::||||||     :::      |||:::
Db      16054 TAGGTTGTCTTTCCCGAATAGCGGGTATCGAGTGTCGGGGTTTGTGGAGGATTCTTACGT 15995

Qy        630 -------------------------PheAlaValPheGlnGluCysGlyLeuTrpValLe 641
                                       |||       |||    |||   ||||||
Db      15994 GCAAGTCGGGTTCAAACACCCTGTTTTCAAAACTCCCCAACTCTGC---CTCTGGTCTGC 15938

Qy        641 uThrAspAlaAsn---LeuThrLysAspTyrIleAspGlyValTyrAsp----------- 656
                 |||          ::::::   ||||||          ||||||
Db      15937 TACCGGTTATAGATATGTGTCTTCGGATTAT---------TATTATGATGGACCGATGAT 15887

Qy        657 ----------------------AsnAlaGluTyrAlaGluArgPhe---MetGluGluAs 668
                                    |||:::      :::   |||    |||||||||
Db      15886 ATTTGTGGGATATGGAGTTCCGGAAGCCAAATCGGTCGATGCTTTCCCGATGGAGGAGAA 15827

Qy        668 nGluGlyHisIleValAspIleHisAspPheSerLeuGlySer----------------- 682
                                ::::::::         |||
Db      15826 AGCTGCTCGGGCAGATGCGGTAAATGAAGCAGAACCCGAATCGAAGGAATCCGGGATTCA 15767

Qy        683 -SerProHisValArgLysHisPheProGluThrTrpIleTrp---------LeuAspTh 699
                     ||||||   :::|||   ||||||      :::          |||   ||
Db      15766 AGAATCCTCCGTGCGTACCAATTTCGCTGAGACTGCATTCTTTGAGCCGGCATTGCTGAC 15707

Qy        699 rAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSe 719
              |:::       :::          |||      |||:::|||:::::::::|||
Db      15706 CGATGAACGGGGCGAAGTTTCTTGGTCTTTC------ACTCTGCCCGAGACACTGACGCG 15653

Qy        719 rTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrPr 739
                |||          |||      ::::::|||:::   ||||||:::
Db      15652 CTGGCACCTGCTCCTCTTTGCACATACGAAGGATATGCGGTTGGGCATGAAAGACGAAAG 15593

Qy        739 oValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIl 759
                ||||||::::::      :::   |||    :::      |||||||||      :::
Db      15592 CGTGGAGGTGCGG------AAAGACTTTATGCTTACACCCAATTTGCCACGATTCCTGAG 15539

Qy        759 eArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGl 779
                  |||::::::                  :::|||   |||       :::          ::
Db      15538 AATGGGTGACAAGGGTACGGCTTCGGCTTCGATTCGGAACGGGAGCGAGACGATGCAGCA 15479

Qy        779 uValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIl 799
              :      |||   :::|||      |||                    ::::::::
Db      15478 GGGCTTCGTCCGTATGGAACTGTTCGAC------------------CCTGCTACCGACAA 15437

Qy        799 eAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLe 819
                    |||   :::            |||    :::          ||||||
Db      15436 GCTATTGGGCGGCGAGAGGCTTCCGTTCTCGGTGGAAGCCGGTGGTACGGTTACGGTGTC 15377

Qy        819 uPheProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerPr 839
                ||||||:::   |||
Db      15376 GTTTCCGCTCGATCCTGTC------------------------------------------ 15358

Qy        839 oThrAlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSe 859
                :::       ||||||:::       :::|||   ::::::
Db      15357 -TCCGGATATGATGCTCTGGGTGTACGTTTGTTAGCCGAAAGCCATGACTTCAGCGATGG 15299

Qy        859 rTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysTh 879
                       ||||||:::               |||:::       |||:::
Db      15298 AGAGCAACACTCGATCGTGCAACTGCCGGCTACGGAGCGAGTGGTGGAGACCATACCGTT 15239

Qy        879 rLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAl 899
                :::    :::               ||||||
Db      15238 GATTCTCTACGGCGGACAGTCCCAAACGGTG----------------------------- 15208

Qy        899 aIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTy 919
                                               |||   ||||||       |||::
Db      15207 -------------------------------GATCTGGATAGTCTA------TTTCCCCA 15185

Qy        919 rGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLe 939
              :      |||          |||          |||      :::   :::   :::   ||
Db      15184 TCGCAGCGGCAGACCGGCTATGGGTACAATGGCACTTCAGGTG---GTGAGCAATCCGCT 15128

Qy        939 uThrLysLysLysGln------------LeuThrAspAsnLeuLysGluLysAlaLeuSe 955
              |          |||                :::||||||      :::|||   |||:::||
Db      15127 TTGGGTGGCCGTACAGGCCTTGCCCGTGATGACTGATGTACACGAGGAAGATGCCGTTTC 15068

Qy        955 rPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSe 975
              |          |||          :::                   ||||||   ::
Db      15067 GGTGGCTTCGGCTCTTTATGCCAATAGTATTGCTGCTGCCCTTGTTTC-GGGTCGCTTGC 15009

Qy        975 rAlaPheGlyAsnTyr------AspProSerGlySerThrTrpLeuSerAlaPheVal-- 992
              :|||:::   :::         :::       |||:::|||||||   |||   ::::::
Db      15008 TGCATACTCCTGCTTTGGGTGAGAGTT-TGCGGCACTACTTGGCGCAGCCCTTGGATACA 14950

Qy        993 ---------------------------------------------------LeuArgCysPh 996
                                                               ||||||::
Db      14949 CTCGCTTTACAAAAATCGCCCTTGTCCTCGGACGAATTGCCTTGGCGCGCGGAGATGCTG 14890

Qy        996 eLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTr 1016
              :   |||   :::                :::          |||             |||||
Db      14889 GCC-GAACAGAACAATCGTCAGCGTTTGCAGGCTCTCGTGGCTTCGGATCGTCTGACATG 14831

Qy       1016 p-----------------------LeuLysGlyHisGlnLysSerAsnGlyGluPhe-- 1027
              |                       ||||||     |||||||||:::|||     :::
```

```
Db      14830 GACGGAGCAGATGCTGACTGACAAACTGAAAAGGCTGCAAAAGTCTGACGGTTCGTGGGC 14771

Qy       1028 -TrpAspProGlyArgValIleHisSerGluLeuGlnGlyGlyAsnLysSerProValTh 1047
               |||                 |||   |||:::          :::
Db      14770 TTGG-----------------CATCCTGAGATGTCTTCCAACGAT------------TA 14741

Qy       1047 rLeuThrAlaTyrIleValThrSerLeuLeu 1057
               ||||||   |||:::::|||    |||:::
Db      14740 TCTGACGGACTATGTGATGACGATGCTGGTA 14710

RESULT 14
US-09-311-352B-1
; Sequence 1, Application US/09311352B
; Patent No. 6329500
; GENERAL INFORMATION:
;  APPLICANT:  Webb, Donna J.
;  APPLICANT:  Gonias, Steven L.
;  TITLE OF INVENTION:  Transforming Growth Factor-beta Binding Site
;  FILE REFERENCE:  00370-02
;  CURRENT APPLICATION NUMBER: US/09/311,352B
;  CURRENT FILING DATE:  1999-05-13
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 339
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-311-352B-1

Alignment Scores:
Pred. No.:              7.98e-16      Length:        339
Score:                  234.00        Matches:       51
Percent Similarity:     58.56%        Conservative:  14
Best Local Similarity:  45.95%        Mismatches:    36
Query Match:            3.18%         Indels:        10
DB:                     4             Gaps:          3

US-10-020-095-4 (1-1428) x US-09-311-352B-1 (1-339)

Qy        660 TyrAlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSer 679
               |||         |||                   :::|||   :::   :::
Db         28 TATGAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG------ 81

Qy        680 LeuGlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrp 696
                          ||||||         ||||||||:::||||||||||||||||||||
Db         82 ------------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGG 129

Qy        697 LeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSer 716
                          :::               |||      ||||||||||||||:::
Db        130 GATTTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACC 189

Qy        717 IleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThr 736
               ||||||   |||   |||         |||   :::|||||||||   ||||||||::::::
Db        190 ATCACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCT 249

Qy        737 ThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyr 756
              :::|||         |||:::||||||||||||||||||:::   |||    :::||||||
Db        250 TCCACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTCACAATGCCTTAC 306

Qy        757 SerValIleArgGlyGluGluPheAlaLeuGlu 767
               ||||||||||||||||||   |||   |||:::
Db        307 TCTGTGATTCGTGGAGAGGCCTTCACACTCAAG 339

RESULT 15
US-09-241-606-3
; Sequence 3, Application US/09241606
; Patent No. 6472140
; GENERAL INFORMATION:
;  APPLICANT: Tanzi, Rudolph E.
;  APPLICANT: Kovacs, Dora
;  APPLICANT: Saunders, Aleister J.
;  TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
;  TITLE OF INVENTION: Alzheimer's Disease
;  FILE REFERENCE: 0609.4460003
;  CURRENT APPLICATION NUMBER: US/09/241,606
;  CURRENT FILING DATE: 1999-02-02
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 750
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(750)
;   OTHER INFORMATION: A /LRP Binding Domain
US-09-241-606-3

Alignment Scores:
Pred. No.:              3.08e-13      Length:        750
Score:                  215.50        Matches:       65
Percent Similarity:     45.08%        Conservative:  45
Best Local Similarity:  26.64%        Mismatches:    95
Query Match:            2.93%         Indels:        39
DB:                     4             Gaps:          7

US-10-020-095-4 (1-1428) x US-09-241-606-3 (1-750)

Qy       1164 IleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAsp 1183
              |||::::::|||:::::::::|||:::|||:::   ||||||||:::||||||||||||
Db         28 ATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAGGGCGGTTTCTCCTCCACCCAGGAC 87

Qy       1184 ThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArgThr 1203
              |||   |||||||||   |||||||||::::::   |||            |||   ||||||
Db         88 ACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGAGCCGCCACATTTACC---AGGACT 144

Qy       1204 AsnIleGlnValThrValThrGlyProSerSerProSerProLeuAlaValValGln--- 1222
                          ||||||      ||||||   :::         :::      |||
Db        145 GGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGGACATTTTCCAGCAAATTCCAAGTG 204

Qy       1223 ------------------------------------------------ProMetAlaVal 1226
                                                               |||
Db        205 GACAACAACAATCGCCTGTTACTGCAGCAGGTCTCATTGCCAGAGCTGCCTGGGGAATAC 264

Qy       1227 AsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnVal 1246
              ::::::         |||   |||               |||   ::::::   ||||||:::
Db        265 AGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTCCAGACCTCCTTGAAATACAATATT 324

Qy       1247 LysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeu 1266
                                               :::      :::|||   |||
Db        325 ----------------------------------CTCCCAGAAAAGGAAGAGTTCCCCTTT 351

Qy       1267 AspValAlaValLys---------GluAsnLysAspAspLeuAsn---HisValAspLeu 1282
               :::   |||:::            :::      |||:::            |||
Db        352 GCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGATGAACCCAAAGCCCACACCAGCTTC 411

Qy       1283 AsnValCysThrSerPheSerGlyProGly---------ArgSerGlyMetAlaLeuMet 1299
                 :::      |||   |||      |||               |||   ||||||:::::::
Db        412 CAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGCTCTGCCTCCAACATGGCGATCGTT 471

Qy       1300 GluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeu--------- 1316
              :::|||   ::::::|||||||||:::                  :::   :::
Db        472 GATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAGCCAACAGTGAAAATGCTTGAAAGA 531

Qy       1317 SerGluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSer 1336
              |||      |||   :::   |||                  :::   :::|||||||||
Db        532 TCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAACCATGTCTTGATTTACCTTGATAAG 591

Qy       1337 ValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsn 1356
              |||:::               :::            ::::::::      |||   :::
Db        592 GTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTTCTGCAAGATGTCCCAGTAAGAGAT 651

Qy       1357 ThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaValArg 1376
                 :::   |||   |||   :::   ||||||||||||            |||:::
Db        652 CTCAAACCAGCCATAGTGAAAGTCTATGATTACTACGAGACGGATGAGTTTGCAATCGCT 711
```

```
Qy       1377 SerTyrAsnSer 1380
                 ||||||:::
Db        712 GAGTACAATGCT 723

Search completed: January 17, 2004, 01:18:53
Job time : 445 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         January 17, 2004, 01:19:00 ; Search time 982 Seconds
                                             (without alignments)
                                             5125.623 Million cell updates/sec

Title:          US-10-020-095-4
Perfect score:  7348
Sequence:       1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters:      4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10020095/runat_15012004_163824_18739/app_query.fasta_1.1607
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10020095_@CGN_1_1_136_@runat_15012004_163824_18739
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
                14:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                17:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                18:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                        %
Result                Query
  No.    Score  Match Length DB  ID                                 Description
------------------------------------------------------------------------------
    1  4535.5   61.7   3033 10  US-09-833-381-1810            Sequence 1810, Ap
    2    3360   45.7   2273 12  US-10-108-260A-953            Sequence 953, App
    3    1441   19.6   4576 13  US-10-292-081A-7              Sequence 7, Appli
    4    1441   19.6   4577 14  US-10-052-817-1               Sequence 1, Appli
    5    1440   19.6   4422 9  US-09-873-403-4                Sequence 4, Appli
    6    1440   19.6   4577 9  US-09-873-403-3                Sequence 3, Appli
    7    1440   19.6   4577 10  US-09-880-107-2236            Sequence 2236, Ap
    8    1440   19.6   4577 12  US-10-331-496A-19             Sequence 19, Appl
    9    1440   19.6   4577 13  US-09-873-319-408             Sequence 408, App
   10    1440   19.6   4577 13  US-09-960-706-654             Sequence 654, App
   11    1440   19.6   4577 13  US-10-292-081A-3              Sequence 3, Appli
   12    1440   19.6   4577 13  US-10-292-081A-5              Sequence 5, Appli
   13    1440   19.6   4577 15  US-10-076-816-56              Sequence 56, Appl
   14    1440   19.6   4809 13  US-10-240-965-178             Sequence 178, App
   15    1438   19.6   4527 13  US-09-756-247-3               Sequence 3, Appli
   16    1438   19.6   4527 13  US-09-756-247-5               Sequence 5, Appli
   17    1429   19.4   4530 13  US-10-292-081A-2              Sequence 2, Appli
   18    1422   19.4   4488 13  US-09-981-151A-9              Sequence 9, Appli
   19    1418   19.3   4823 13  US-09-971-392-172             Sequence 172, App
   20  1412.5   19.2   4595 10  US-09-917-800A-1531           Sequence 1531, Ap
   21    1410   19.2   5092 13  US-10-006-285-404             Sequence 404, App
   22    1410   19.2   5092 13  US-09-971-429B-2              Sequence 2, Appli
   23    1400   19.1   4615 10  US-09-880-107-3727            Sequence 3727, Ap
   24    1390   18.9   4615 13  US-10-006-285-405             Sequence 405, App
   25    1384   18.8   4677 13  US-10-316-253-266             Sequence 266, App
   26    1384   18.8   4701 13  US-10-316-253-268             Sequence 268, App
   27    1312   17.9   1300 13  US-10-133-013-223             Sequence 223, App
   28    1274   17.3   4771 13  US-09-756-247-27              Sequence 27, Appl
   29    1255   17.1   4021 9  US-09-925-301-552              Sequence 552, App
   30    1157   15.7   4426 12  US-10-094-886-123             Sequence 123, App
   31  1109.5   15.1   4501 12  US-10-094-886-125             Sequence 125, App
   32  1072.5   14.6   5191 13  US-10-006-285-206             Sequence 206, App
   33   991.5   13.5   5555 15  US-10-198-846-13979           Sequence 13979, A
   34     988   13.4   5417 13  US-10-007-926A-413            Sequence 413, App
   35     956   13.0   5087 15  US-10-001-076-111             Sequence 111, App
   36     934   12.7   5067 10  US-09-964-824A-546            Sequence 546, App
   37     934   12.7   5067 11  US-09-918-624B-18             Sequence 18, Appl
   38     934   12.7   5067 13  US-10-101-510-316             Sequence 316, App
   39     934   12.7   5067 15  US-10-001-076-3               Sequence 3, Appli
   40     931   12.7   5067 9  US-09-875-519A-23              Sequence 23, Appl
   41     930   12.7    875 13  US-10-160-162-48              Sequence 48, Appl
   42     930   12.7    875 13  US-09-820-649-48              Sequence 48, Appl
   43     930   12.7   5090 14  US-10-044-090-470             Sequence 470, App
   44   925.5   12.6   6435 15  US-10-001-076-10              Sequence 10, Appl
   45   921.5   12.5   4871 9  US-09-778-927A-18              Sequence 18, Appl

                                 ALIGNMENTS

RESULT 1
US-09-833-381-1810
; Sequence 1810, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1810
;   LENGTH: 3033
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-1810

Alignment Scores:
Pred. No.:              0             Length:         3033
Score:                  4535.50       Matches:        880
Percent Similarity:     98.00%        Conservative:   2
```

```
Best Local Similarity:  97.78%       Mismatches:    1
Query Match:            61.72%       Indels:       17
DB:                     10           Gaps:          1

US-10-020-095-4 (1-1428) x US-09-833-381-1810 (1-3033)

Qy        546 ValLeuLysIleProValGlnLeuValPheLysAsnLysIleLysLeuTyrTrpSerLys 565
              |||   |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 GTCCGAAAAATTCCTGTTCAGCTTGTTTTAAAAATAAGATAAAGCTATATTGGAGTAAA 65

Qy        566 ValLysAlaGluProSerGluLysValSerLeuArgIleSerValThrGlnProAspSer 585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTGACACAGCCTGACTCC 125

Qy        586 IleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIle 605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 ATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAATGCCTCTAATGATATT 185

Qy        606 ThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMet 625
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 ACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGATATTATTTAGGCATG 245

Qy        626 PheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsn 645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTATTGACAGATGCAAAC 305

Qy        646 LeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyrAlaGluArgPheMet 665
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATATGCTGAGAGGTTTATG 365

Qy        666 GluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeuGlySerSerProHis 685
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTGGGTAGCAGTCCACAT 425

Qy        686 ValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIle 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 GTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATGGGTTACAGGATT 485

Qy        706 TyrGlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPhe 725
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 TACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTT 545

Qy        726 ValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPhe 745
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 GTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTC 605

Qy        746 GlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAla 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 CAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGAATTTGCT 665

Qy        766 LeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIleGlu 785
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 TTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAGGTAATCATTGAG 725

Qy        786 LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGln 805
              ||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db        726 AAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAATGCCACAGGCCACCAG 785

Qy        806 GlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgProThr 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 CAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTTCCCATCAGGCCAACA 845

Qy        826 HisLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAlaSerAspAlaVal 845
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        846 CATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACTGCTTCTGATGCTATC 905

Qy        846 ThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeu 865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 ACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATATTCACAATCCATCTTA 965

Qy        866 LeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPhePro 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        966 TTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGTTTCTCATTTCCT 1025

Qy        886 ProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeuGly 905
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 CCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGAGATGTTCTTGGT 1085

Qy        906 ProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsn 925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 CCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGTGGTGAACAGAAC 1145

Qy        926 MetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeu 945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 ATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAAAAGAAACAACTG 1205

Qy        946 ThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeu 965
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 ACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTACCAGAGAGAACTT 1265

Qy        966 LeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySer 985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 CTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGACCCTTCTGGGAGC 1325

Qy        986 ThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAspIle 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 ACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCTTACATAGATATT 1385

Qy       1006 AspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGly 1025
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1386 GATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAGAAATCCAACGGT 1445

Qy       1026 GluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGlyAsnLysSerPro 1045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 GAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCA 1505

Qy       1046 ValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsn 1065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 GTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAAC 1565

Qy       1066 IleAspValGlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAsp 1085
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 ATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGAC 1625

Qy       1086 AsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLys 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 AATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAG 1685

Qy       1106 GluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpVal 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1686 GAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTG 1745

Qy       1126 SerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAla 1145
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1746 TCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCA 1805

Qy       1146 AlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMet 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1806 GCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATG 1865

Qy       1166 ArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThr 1185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1866 AGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACT 1925

Qy       1186 ValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIle 1205
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1926 GTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATC 1985

Qy       1206 GlnValThrValThrGlyProSerSerProSerPro------------------------ 1217
              ||||||||||||||||||||||||||||||||||||
Db       1986 CAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTTCTGATTGACACACAC 2045

Qy       1218 ---------------------------LeuAlaValValGlnProMetAlaValAsnIle 1228
                                         |||||||||||||||||||||||||||||||||
Db       2046 AACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCAATGGCAGTTAATATT 2105
```

```
Qy       1229 SerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAla 1248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2106 TCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCT 2165

Qy       1249 SerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspVal 1268
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2166 TCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTT 2225

Qy       1269 AlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPhe 1288
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2226 GCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTT 2285

Qy       1289 SerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMet 1308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2286 TCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATG 2345

Qy       1309 ValProSerGluAlaIleSerLeuSerGluThrValLysLysValGluTyrAspHisGly 1328
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2346 GTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGA 2405

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2406 AAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCT 2465

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2466 GTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTAT 2525

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCysAsp 1388
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2526 GAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGAC 2585

Qy       1389 LeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAlaSerGlySerHisHis 1408
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2586 CTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCAT 2645

Qy       1409 HisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPheMetGluLeuTrpLeu 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2646 CACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTG 2705
```

RESULT 2
US-10-108-260A-953
; Sequence 953, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 953
;   LENGTH: 2273
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-108-260A-953

```
Alignment Scores:
Pred. No.:               0             Length:         2273
Score:                   3360.00       Matches:        657
Percent Similarity:      100.00%       Conservative:   0
Best Local Similarity:   100.00%       Mismatches:     0
Query Match:             45.73%        Indels:         0
DB:                      12            Gaps:           0
```

US-10-020-095-4 (1-1428) x US-10-108-260A-953 (1-2273)

```
Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 125

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 185

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 245

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 305

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 365

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 425

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 485

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 545

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 605

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 665

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAG 725

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 785

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 845

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        846 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 905

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 965

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        966 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1025

Qy        321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1085

Qy        341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1145

Qy        361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1205

Qy        381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1206 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1265

Qy        401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1325

Qy        421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1385

Qy        441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1386 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1445

Qy        461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1505

Qy        481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1565

Qy        501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1625

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1685

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1686 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1745

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1746 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1805

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1806 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1865

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1866 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1925

Qy        621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1926 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1985

Qy        641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsn 657
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1986 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAAT 2036
```

RESULT 3
US-10-292-081A-7
; Sequence 7, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLO
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-292-081A-7

```
Alignment Scores:
Pred. No.:                2.83e-133     Length:         4576
Score:                    1441.00       Matches:        453
Percent Similarity:       46.20%        Conservative:   264
Best Local Similarity:    29.19%        Mismatches:     551
Query Match:              19.61%        Indels:         284
DB:                       13            Gaps:           60
```

US-10-020-095-4 (1-1428) x US-10-292-081A-7 (1-4576)

```
Qy         12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
              |||||    :::                      :::|||:::         |||:::::::::|||
Db         85 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 144

Qy         32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
              |||   :::::                      |||      ||||||        :::
Db        145 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 201

Qy         52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
                       ||||||||   |||   |||                  |||   ::::::
Db        202 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 261

Qy         72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
                 :::   |||||||||||      |||             :::|||
Db        262 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 309

Qy         90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
                    ::::::|||   :::|||::::::|||                  ::::::|||
Db        310 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 363

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
              |||   ||||||:::                   |||         :::      :::
Db        364 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 414

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
              :::         ||||||:::||||||||||||::::::|||||||||   |||   |||
Db        415 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 474

Qy        147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
              ||||||||:::|||::::::         :::|||   |||               :::   :::
Db        475 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 534

Qy        167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
              :::   |||:::||||||||   |||   |||   ||||||   |||   |||   :::
Db        535 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 594

Qy        186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
                 |||:::            :::|||   |||||||||   |||      |||   :::
Db        595 GGTGGCCTCAAGCAATTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 654

Qy        206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
              :::   ||||||                  :::|||   :::         |||   |||   |||
Db        655 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 711

Qy        224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
              :::||||||||||||||||||||||||   :::         |||         ::::::
Db        712 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 771

Qy        244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
              :::   :::|||   ::::::            ||||||||||||||||||   |||
Db        772 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGGAAGCCTGTCCCTGGACAT 831

Qy        264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
              ||||||::::::
Db        832 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 891

Qy        272 ------------------------------------------SerPheTrpGlyLys 276
```

```
Db    892 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 951

Qy    277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296

Db    952 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG-----------------AAGGAG 987

Qy    297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316

Db    988 TATGAAATGAAA-----------------------CTTCACACTGAGGCCCAGATCCAA 1023

Qy    317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331

Db   1024 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1083

Qy    332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351

Db   1084 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1131

Qy    352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371

Db   1132 CCCTTCTTTGGG---------------------------------------------CAG 1146

Qy    372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391

Db   1147 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1194

Qy    392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411

Db   1195 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1245

Qy    412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428

Db   1246 CATGGCCTTGTACAGTTCTCTATCAACACCACCAATGTTATGGGTACCTCTCTTACTGTT 1305

Qy    429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438

Db   1306 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1365

Qy    439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458

Db   1366 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1389

Qy    459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478

Db   1390 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1449

Qy    479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492

Db   1450 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1509

Qy    493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510

Db   1510 CTGAAGAAGCTCTCCTTCTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1569

Qy    511 ------------------LysGlnAsnSerThr------------MetPheSerLeuThr 520

Db   1570 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1629

Qy    521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540

Db   1630 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1689

Qy    541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560

Db   1690 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1749

Qy    561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580

Db   1750 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1809

Qy    581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMet--- 599

Db   1810 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1866

Qy    600 ---------------AsnAlaSer------------------------------------ 602

Db   1867 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1926

Qy    603 ------------------AsnAspIleThrMetGluAsnValValHisGluLeuGluLeu 616

Db   1927 GGCTTCCCTGGGCCTTTGAATGACCAGGACAATGAAGACTGCATCAATCGTCATAATGTC 1986

Qy    617 TyrAsnThrGlyTyr---TyrLeuGlyMetPheMetAsnSer------------PheAla 631

Db   1987 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2046

Qy    632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651

Db   2047 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2106

Qy    652 AspGlyValTyrAspAsnAlaGlu--------------------------------Tyr 660

Db   2107 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2166

Qy    661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680

Db   2167 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2217

Qy    681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697

Db   2218 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2268

Qy    698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717

Db   2269 TTGGTGGTGGTAAACTCAGCAGGTGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2328

Qy    718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737

Db   2329 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2388

Qy    738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757

Db   2389 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTCACAATGCCTTACTCT 2445

Qy    758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777

Db   2446 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2505

Qy    778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797

Db   2506 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTAGCTGTCCCAGTGGAGAAG 2565

Qy    798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812

Db   2566 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA-------------------- 2604

Qy    813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832

Db   2605 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2652

Qy    833 ThrValThrAlaLeuSer---------------------------------ProThrAla 841

Db   2653 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2712

Qy    842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856

Db   2713 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2772

Qy    857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876

Db   2773 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2826

Qy    877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896

Db   2827 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2883
```

```
Qy        897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
              ::::::   :::|||||:::|||||    ::::::             :::|||::::::
Db       2884 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2943

Qy        917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
              |||||||||||||||||||||||||||||:::   ||||||||||||||||:::|||
Db       2944 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3003

Qy        937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
              ||||||||   :::   :::||||||||        :::|||   ||||||:::   :::
Db       3004 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGATCAAGTCCAAGGCCATTGGCTAT 3063

Qy        957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
              :::      ||||||||||||:::|||   |||:::         |||||||||:::|||
Db       3064 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3123

Qy        977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
              ||||||      |||      :::      |||:::|||||||||:::||||||||||||:::
Db       3124 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3183

Qy        995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
                 |||   :::|||      ||||||   ||||||:::         :::   :::
Db       3184 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3243

Qy       1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
                 ||||||         ||||||   ||||||   |||         |||   ::::::
Db       3244 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3303

Qy       1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
              ::::::   ::::::||||||   :::         |||||||||:::|||||||||
Db       3304 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3363

Qy       1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
              :::||||||                     :::      ::::::
Db       3364 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3423

Qy       1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
              |||||||||                  |||         |||   |||      ||||||
Db       3424 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3480

Qy       1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
                 ||||||:::   ||||||   :::      |||:::      |||   ||||||
Db       3481 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3540

Qy       1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
              |||      |||         |||   ::::::            |||
Db       3541 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3600

Qy       1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
                 ::::::   :::      ::::::|||:::            |||   ::::::
Db       3601 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3660

Qy       1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
              |||:::   :::|||   ||||||:::   :::|||         |||      ||||||
Db       3661 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3717

Qy       1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
              |||   :::               |||::::::|||:::::::::|||:::|||:::
Db       3718 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3777

Qy       1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
              |||||||||:::|||||||||||||||   |||||||||   |||||||||::::::
Db       3778 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3837

Qy       1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
              |||      |||   ||||||               ||||||      ||||||
Db       3838 GCAGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3894

Qy       1216 SerProLeuAlaValValGln--------------------------------------- 1222
              :::      :::   |||
Db       3895 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAACCGCCTGTTACTGCAGCAGGTCTCA 3954

Qy       1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
                          |||         ::::::               |||   |||
Db       3955 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4014

Qy       1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
              |||   ::::::   ||||||:::                                 :::
Db       4015 CAGACATCCTTGAAATACAATATT---------------------------------CTC 4041

Qy       1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
                 :::|||   |||         :::   |||:::         :::      |||
Db       4042 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4101

Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
              :::            |||         :::      |||   |||      |||
Db       4102 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4161

Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
                 |||   ||||||:::::::::|||   ::::::|||||||||:::
Db       4162 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4221

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
              :::   :::                  |||      |||   :::   |||
Db       4222 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4281

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
                 :::   :::||||||||   |||:::                  :::
Db       4282 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4341

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
              :::::::::      |||   :::   :::   |||   |||   :::   ||||||||
Db       4342 CTGCAAGATGTCCCAGTAAGAGATCTGAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4401

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
              |||         |||:::      ||||||:::
Db       4402 GAGACGGATGAGTTTGCAATTGCTGAGTACAATGCT 4437
```

RESULT 4
US-10-052-817-1
; Sequence 1, Application US/10052817
; Publication No. US20020114792A1
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Kovacs, Dora
; APPLICANT: Saunders, Aleister J.
; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 0609.4460005
; CURRENT APPLICATION NUMBER: US/10/052,817
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: 09/241,606
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: 09/148,503
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/093,297
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4577
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (44)..(112)
; NAME/KEY: CDS
; LOCATION: (44)..(4468)
; NAME/KEY: mat_peptide
; LOCATION: (113)..(4468)
US-10-052-817-1

Alignment Scores:
Pred. No.: 2.83e-133 Length: 4577
Score: 1441.00 Matches: 451

```
Percent Similarity:       46.20%      Conservative:   266
Best Local Similarity:    29.06%      Mismatches:     551
Query Match:              19.61%      Indels:         284
DB:                       14          Gaps:           59

US-10-020-095-4 (1-1428) x US-10-052-817-1 (1-4577)

Qy      12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db      86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy      32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
Db     146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy      52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
Db     203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy      72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
Db     263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy      90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
Db     311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy     107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db     365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy     127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
Db     416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy     147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db     476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy     167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
Db     536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy     186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db     596 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy     206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db     656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy     224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db     713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772

Qy     244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db     773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGAAGCCTGTCCCTGGACAT 832

Qy     264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
Db     833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy     272 ------------------------------------------------SerPheTrpGlyLys 276
Db     893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952

Qy     277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db     953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 988

Qy     297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
Db     989 TATGAAATGAAA------------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy     317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
Db    1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy     332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db    1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy     352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
Db    1133 CCCTTCTTTGGG---------------------------------------------CAG 1147

Qy     372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
Db    1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1195

Qy     392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
Db    1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy     412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
Db    1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy     429 LysIleGluPhePro----------------------------IleLeuGluAspSer 438
Db    1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy     439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db    1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390

Qy     459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db    1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy     479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492
Db    1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy     493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
Db    1511 CTGAAGAAGCTCTCCTTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy     511 ------------------LysGlnAsnSerThr------------MetPheSerLeuThr 520
Db    1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy     521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db    1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1690

Qy     541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db    1691 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750

Qy     561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db    1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810

Qy     581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db    1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867

Qy     601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
Db    1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927

Qy     620 GlyTyr------------------------------------------TyrLeuGlyMet 625
Db    1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987
```

```
Qy     626 PheMetAsnSer----------------------------------------PheAla 631
           :::::|||                                        ::::::
Db    1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047

Qy     632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
              ::::::  ||||||          |||::::::  :::  |||       :::
Db    2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107

Qy     652 AspGlyValTyrAspAsnAlaGlu--------------------------------Tyr 660
                               |||                                |||
Db    2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167

Qy     661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
                           |||                   :::|||  :::   :::
Db    2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218

Qy     681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
                    ||||||         ||||||||:::|||||||||||||||||||
Db    2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269

Qy     698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
              :::                           |||     ||||||||||||||:::|||
Db    2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329

Qy     718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
           |||   |||   |||         |||   :::|||||||||      ||||||||:::::::::
Db    2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389

Qy     738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
           |||         |||:::|||||||||||||||||:::   |||   :::|||||||||
Db    2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446

Qy     758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
           ||||||||||||||||||   |||   |||:::   |||:::   |||||||||
Db    2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506

Qy     778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
              |||   |||   :::|||   |||       |||      :::
Db    2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566

Qy     798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
           |||               |||   |||   |||   |||
Db    2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA--------------------- 2605

Qy     813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
                       ||||||   :::   :::   |||      ||||||   :::
Db    2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy     833 ThrValThrAlaLeuSer--------------------------------ProThrAla 841
           ||||||:::|||   :::                                |||:::
Db    2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy     842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
                          |||   |||   :::   :::||||||:::   ||||||:::
Db    2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy     857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
           ||||||      :::   :::   ||||||         :::         :::
Db    2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy     877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
              :::   ||||||         |||||||||   |||         |||   |||
Db    2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy     897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
           ::::::   :::||||||:::||||||         :::||:::::
Db    2885 GTCTCAGTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy     917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
           ||||||||||||||||||||||||||||||:::   ||||||||||||||||||:::|||
Db    2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy     937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
           |||||||||   :::   :::|||||||||      :::|||   ||||||:::   :::
Db    3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGATCAAGTCCAAGGCCATTGGCTAT 3064

Qy     957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
           :::      |||||||||||||:::|||   |||:::      |||||||||:::|||
Db    3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy     977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
           ||||||   |||      :::   |||:::|||||||||:::||||||||||||:::
Db    3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy     995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
              |||   :::|||      ||||||   ||||||:::         :::   :::
Db    3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244

Qy    1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
              ||||||         ||||||   ||||||   |||         |||   ::::::
Db    3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy    1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
           ::::::   ::::::||||||   :::         ||||||||||:::|||||||||
Db    3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy    1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
           :::||||||                     :::         |||:::   ::::::
Db    3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy    1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
           |||||||||                  |||      |||   |||      ||||||
Db    3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy    1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
              ||||||:::   ||||||      :::      |||:::      |||   ||||||
Db    3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy    1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
           |||   |||      |||   ::::::                  |||
Db    3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy    1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
              ::::::   :::         ::::::|||:::         |||   ::::::
Db    3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy    1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
           |||:::   :::|||   ||||||:::   :::|||      |||         ||||||
Db    3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy    1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
           |||   :::               |||::::::|||:::::::::|||:::|||:::
Db    3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy    1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
           ||||||||:::||||||||||||||||||   ||||||   ||||||||||:::::
Db    3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy    1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
           |||         |||   ||||||            ||||||      ||||||
Db    3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy    1216 SerProLeuAlaValValGln--------------------------------------- 1222
           :::      :::      |||
Db    3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy    1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
                       |||      ::::::            |||   |||
Db    3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy    1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
           |||   ::::::   |||||||||:::                              :::
Db    4016 CAGACCTCCTTGAAATACAATATT--------------------------------CTC 4042

Qy    1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
```

```
Db        4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102

Qy        1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292

Db        4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy        1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311

Db        4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy        1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328

Db        4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy        1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348

Db        4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy        1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368

Db        4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy        1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380

Db        4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438

RESULT 5
US-09-873-403-4
; Sequence 4, Application US/09873403
; Patent No. US20020028207A1
; GENERAL INFORMATION:
;  APPLICANT: Srivstava, Pramod K
;  TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC
;  TITLE OF INVENTION:  MOLECULES FOR IMMUNOTHERAPY
;  FILE REFERENCE: 8449-178
;  CURRENT APPLICATION NUMBER: US/09/873,403
;  CURRENT FILING DATE:  2001-06-04
;  PRIOR APPLICATION NUMBER: 09/625,139
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: 60/209,266
;  PRIOR FILING DATE: 2000-06-02
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 4422
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-403-4

Alignment Scores:
Pred. No.:               3.36e-133     Length:        4422
Score:                   1440.00       Matches:       451
Percent Similarity:      46.20%        Conservative:  266
Best Local Similarity:   29.06%        Mismatches:    551
Query Match:             19.60%        Indels:        284
DB:                      9             Gaps:          59

US-10-020-095-4 (1-1428) x US-09-873-403-4 (1-4422)

Qy          12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31

Db          43 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 102

Qy          32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51

Db         103 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 159

Qy          52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71

Db         160 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 219

Qy          72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89

Db         220 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 267

Qy          90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106

Db         268 -----GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 321

Qy         107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126

Db         322 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 372

Qy         127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146

Db         373 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 432

Qy         147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166

Db         433 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 492

Qy         167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185

Db         493 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 552

Qy         186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205

Db         553 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 612

Qy         206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223

Db         613 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 669

Qy         224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243

Db         670 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 729

Qy         244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263

Db         730 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGAAAGCCTGTCCCTGGACAT 789

Qy         264 ValThrLeuThrPheLeuProLeu------------------------------------ 271

Db         790 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 849

Qy         272 ---------------------------------------------SerPheTrpGlyLys 276

Db         850 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 909

Qy         277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296

Db         910 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 945

Qy         297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316

Db         946 TATGAAATGAAA------------------------CTTCACACTGAGGCCCAGATCCAA 981

Qy         317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331

Db         982 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1041

Qy         332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351

Db        1042 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1089

Qy         352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371

Db        1090 CCCTTCTTTGGG---------------------------------------------CAG 1104

Qy         372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391

Db        1105 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1152

Qy         392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
```

```
Db     1153 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1203

Qy      412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
                     ||||||      ||||||   |||        :::       |||   |||
Db     1204 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1263

Qy      429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438
            ::::::   :::                                 :::    |||:::
Db     1264 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1323

Qy      439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
               |||            ||||||
Db     1324 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1347

Qy      459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
             |||     ||||||||||||:::::::::   |||:::                :::
Db     1348 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1407

Qy      479 ValGlySer-----------ProPheGluLeuValValSerGlyAsnLys------Arg 492
               |||                            :::::::::|||
Db     1408 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1467

Qy      493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
            ||||||:::||||||       |||::::::::::::::::|||    :::|||    |||
Db     1468 CTGAAGAAGCTCTCCTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1527

Qy      511 -----------------LysGlnAsnSerThr-----------MetPheSerLeuThr 520
                             ||||||                        |||:::
Db     1528 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1587

Qy      521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
               ::::::         |||   |||      ::::::::|||    :::      |||
Db     1588 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1647

Qy      541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
            ::::::|||   |||         |||      |||:::         ||||||:::
Db     1648 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1707

Qy      561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
            |||   :::|||      ::::::   |||:::            ||||||::::::
Db     1708 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1767

Qy      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
             |||   |||:::      :::   |||||||||:::||||||   ||||||
Db     1768 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1824

Qy      601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
               :::::::::         :::|||    :::   |||   |||          |||
Db     1825 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1884

Qy      620 GlyTyr------------------------------------------TyrLeuGlyMet 625
            |||:::                                                   :::
Db     1885 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1944

Qy      626 PheMetAsnSer------------------------------------------PheAla 631
            ::::::|||                                             ::::::
Db     1945 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2004

Qy      632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
               ::::::   ||||||         |||::::::   :::   |||         :::
Db     2005 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2064

Qy      652 AspGlyValTyrAspAsnAlaGlu--------------------------------Tyr 660
                                |||                                  |||
Db     2065 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2124

Qy      661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
                         |||                    :::|||   :::   :::
Db     2125 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2175

Qy      681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
                     ||||||          ||||||||:::||||||||||||||||||
Db     2176 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2226

Qy      698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
                   :::                     |||      ||||||||||||||:::|||
Db     2227 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2286

Qy      718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
            |||   |||   |||         |||   :::|||||||||   |||||||||:::::::::
Db     2287 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2346

Qy      738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
            |||         |||:::||||||||||||||||||:::      |||    :::|||||||||
Db     2347 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2403

Qy      758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
            |||||||||||||||||      |||   |||:::      |||:::   |||||||||
Db     2404 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2463

Qy      778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
                     |||   |||   :::|||   |||         |||      :::
Db     2464 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2523

Qy      798 Glu--------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
            |||                    |||   |||   |||   |||
Db     2524 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA-------------------- 2562

Qy      813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
                        ||||||   :::   :::   |||      ||||||   :::
Db     2563 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2610

Qy      833 ThrValThrAlaLeuSer--------------------------------ProThrAla 841
            ||||||:::|||   :::                                  |||:::
Db     2611 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2670

Qy      842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
                              |||   |||   :::   :::||||||:::   ||||||:::
Db     2671 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2730

Qy      857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
            ||||||      :::   :::   ||||||         :::            :::
Db     2731 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2784

Qy      877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
               :::   ||||||         |||||||||   |||      |||   |||
Db     2785 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2841

Qy      897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
            ::::::   :::||||||:::||||||   ::::::               :::|||::::::
Db     2842 GTCTCAGTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2901

Qy      917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
            |||||||||||||||||||||||||||:::   ||||||||||||||||||:::|||
Db     2902 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 2961

Qy      937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
            |||||||||   :::   :::|||||||||         :::|||   ||||||:::   :::
Db     2962 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3021

Qy      957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
            :::      ||||||||||||:::|||   |||:::      ||||||||||:::|||
Db     3022 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3081

Qy      977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
            ||||||     |||         :::   |||:::|||||||||:::||||||||||||:::
Db     3082 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3141

Qy      995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
             |||   :::|||      ||||||   ||||||:::      :::   :::
Db     3142 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3201

Qy     1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
               ||||||         ||||||   ||||||   |||         |||   ::::::
Db     3202 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3261
```

```
Qy       1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
              :::::: :::::||||||  :::        |||||||||:::|||||||||
Db       3262 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3321

Qy       1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
              :::||||||                   :::      |||:::   ::::::
Db       3322 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3381

Qy       1075 LeuGluSer-----------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
              |||||||||                 |||       |||  |||       ||||||
Db       3382 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3438

Qy       1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
                    ||||||:::  ||||||  :::     |||:::      |||  ||||||
Db       3439 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3498

Qy       1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
              |||   |||          |||  ::::::            |||
Db       3499 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3558

Qy       1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
                 ::::::  :::          ::::::|||:::       |||  ::::::
Db       3559 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3618

Qy       1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
              |||:::   :::|||   ||||||:::   :::|||      |||          ||||||
Db       3619 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3675

Qy       1160 GluGlyIlePro-----------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
              |||    :::              |||::::::|||:::::::::|||:::|||:::
Db       3676 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3735

Qy       1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
              |||||||||:::||||||||||||||||   |||||||||   |||||||||::::::
Db       3736 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3795

Qy       1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
              |||       |||   ||||||                ||||||     ||||||
Db       3796 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3852

Qy       1216 SerProLeuAlaValValGln--------------------------------------- 1222
              :::   :::   |||
Db       3853 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3912

Qy       1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
                          |||      ::::::         |||  |||
Db       3913 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 3972

Qy       1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
              |||   ::::::   ||||||:::                                  :::
Db       3973 CAGACCTCCTTGAAATACAATATT---------------------------------CTC 3999

Qy       1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
                 :::|||  |||       :::        |||:::          :::      |||
Db       4000 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4059

Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
              :::           |||         :::      |||  |||     |||
Db       4060 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4119

Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
                    |||   ||||||:::::::::|||   ::::::|||||||||:::
Db       4120 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4179

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
                 :::   :::            |||      |||   :::   |||
Db       4180 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4239

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
                 :::   :::|||||||||  |||:::                :::
Db       4240 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4299

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
              :::::::::          |||   :::   :::   |||   |||   :::   |||||||||
Db       4300 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4359

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
                 |||        |||:::   ||||||:::
Db       4360 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4395
```

RESULT 6
US-09-873-403-3
; Sequence 3, Application US/09873403
; Patent No. US20020028207A1
; GENERAL INFORMATION:
; APPLICANT: Srivstava, Pramod K
; TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC
; TITLE OF INVENTION: MOLECULES FOR IMMUNOTHERAPY
; FILE REFERENCE: 8449-178
; CURRENT APPLICATION NUMBER: US/09/873,403
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/625,139
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/209,266
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 4577
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-403-3

```
Alignment Scores:
Pred. No.:                3.56e-133    Length:         4577
Score:                    1440.00      Matches:        451
Percent Similarity:       46.20%       Conservative:   266
Best Local Similarity:    29.06%       Mismatches:     551
Query Match:              19.60%       Indels:         284
DB:                       9            Gaps:           59

US-10-020-095-4 (1-1428) x US-09-873-403-3 (1-4577)

Qy         12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
              ||||||   :::         :::|||:::         |||:::::::::|||
Db         86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy         32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
                 |||   ::::::                        |||   ||||||   :::
Db        146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy         52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
                       |||||||||   |||   |||                  |||   ::::::
Db        203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy         72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
                 :::   ||||||||||||      |||            :::|||
Db        263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy         90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
                    ::::::|||   :::|||::::::|||             ::::::|||
Db        311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
                 |||   ||||||:::            |||      |||   :::      :::
Db        365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
              :::            ||||||:::||||||||||||:::::: |||||||||   |||   |||
Db        416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy        147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
              |||||||||:::|||::::::         :::|||   |||              :::   :::
Db        476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy        167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
```

```
Db    536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595
Qy    186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db    596 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655
Qy    206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db    656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712
Qy    224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db    713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772
Qy    244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db    773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGAAGCCTGTCCCTGGACAT 832
Qy    264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
Db    833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892
Qy    272 ---------------------------------------------SerPheTrpGlyLys 276
Db    893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952
Qy    277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db    953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 988
Qy    297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
Db    989 TATGAAATGAAA-----------------------CTTCACACTGAGGCCCAGATCCAA 1024
Qy    317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
Db   1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084
Qy    332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db   1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132
Qy    352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
Db   1133 CCCTTCTTTGGG------------------------------------------CAG 1147
Qy    372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
Db   1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA-----------AATAAAGTCATA 1195
Qy    392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
Db   1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246
Qy    412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
Db   1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306
Qy    429 LysIleGluPhePro------------------------------IleLeuGluAspSer 438
Db   1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366
Qy    439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db   1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390
Qy    459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db   1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450
Qy    479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492
Db   1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510
Qy    493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
Db   1511 CTGAAGAAGCTCTCCTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570
Qy    511 ------------------LysGlnAsnSerThr------------MetPheSerLeuThr 520
Db   1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630
Qy    521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db   1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1690
Qy    541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db   1691 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750
Qy    561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db   1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810
Qy    581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db   1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867
Qy    601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
Db   1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927
Qy    620 GlyTyr------------------------------------------TyrLeuGlyMet 625
Db   1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987
Qy    626 PheMetAsnSer------------------------------------------PheAla 631
Db   1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047
Qy    632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
Db   2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107
Qy    652 AspGlyValTyrAspAsnAlaGlu--------------------------------Tyr 660
Db   2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167
Qy    661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db   2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218
Qy    681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
Db   2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269
Qy    698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
Db   2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329
Qy    718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
Db   2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389
Qy    738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
Db   2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446
Qy    758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
Db   2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506
Qy    778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
Db   2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566
```

```
Qy   798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db  2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA-------------------- 2605

Qy   813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db  2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy   833 ThrValThrAlaLeuSer---------------------------------ProThrAla 841
Db  2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy   842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
Db  2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy   857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
Db  2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy   877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
Db  2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy   897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
Db  2885 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy   917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
Db  2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy   937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
Db  3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3064

Qy   957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
Db  3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy   977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
Db  3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy   995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
Db  3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244

Qy  1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
Db  3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy  1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db  3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy  1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
Db  3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy  1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db  3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy  1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
Db  3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy  1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
Db  3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy  1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
Db  3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy  1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
Db  3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy  1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
Db  3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy  1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
Db  3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy  1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
Db  3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy  1216 SerProLeuAlaValValGln--------------------------------------- 1222
Db  3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy  1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
Db  3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy  1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
Db  4016 CAGACCTCCTTGAAATACAATATT--------------------------------CTC 4042

Qy  1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
Db  4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102

Qy  1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
Db  4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy  1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
Db  4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy  1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
Db  4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy  1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
Db  4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy  1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
Db  4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy  1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
Db  4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438
```

RESULT 7
US-09-880-107-2236
; Sequence 2236, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14

```
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2236
;   LENGTH: 4577
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 M11313
US-09-880-107-2236

Alignment Scores:
Pred. No.:              3.56e-133      Length:        4577
Score:                  1440.00        Matches:       451
Percent Similarity:     46.20%         Conservative:  266
Best Local Similarity:  29.06%         Mismatches:    551
Query Match:            19.60%         Indels:        284
DB:                     10             Gaps:          59

US-10-020-095-4 (1-1428) x US-09-880-107-2236 (1-4577)

Qy          12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
               ||||||   :::              :::|||:::       |||::::::::::|||
Db          86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy          32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
                  |||   ::::::                          |||       ||||||   :::
Db         146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy          52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
                           |||||||||   |||   |||                |||   ::::::
Db         203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy          72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
                  :::   |||||||||||||               :::|||
Db         263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy          90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
                     ::::::|||   :::|||::::::|||                  ::::::|||
Db         311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy         107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
                  |||   ||||||:::            |||       |||   :::      :::
Db         365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy         127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
               :::         ||||||:::||||||||||||::::::|||||||||   |||   |||
Db         416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy         147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
               |||||||||:::|||::::::               :::|||         :::   :::
Db         476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy         167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
               :::   |||:::|||||||||   |||   |||   ||||||   |||   |||   :::
Db         536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy         186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
                  |||:::         :::|||   |||||||||   |||      |||   :::
Db         596 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy         206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
               :::   ||||||            :::|||   :::      |||   |||   |||
Db         656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy         224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
               :::||||||||||||||||||||||||   :::      |||         ::::::
Db         713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772

Qy         244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
               :::   :::|||   ::::::               ||||||||||||||||||   |||
Db         773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGGAAGCCTGTCCCTGGACAT 832

Qy         264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
               ||||||::::::
Db         833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy         272 ------------------------------------------SerPheTrpGlyLys 276
                                                         |||:::   :::
Db         893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952

Qy         277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
                  |||      ||||||   |||::::::                           :::
Db         953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG-----------------AAGGAG 988

Qy         297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
                  |||||||||                       |||                :::
Db         989 TATGAAATGAAA-----------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy         317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
                     |||   |||      ||||||                   :::|||   ::::::
Db        1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy         332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
               |||   :::|||      |||         |||:::|||                  |||
Db        1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy         352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
                  ||||||                                                :::
Db        1133 CCCTTCTTTGGG------------------------------------------CAG 1147

Qy         372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
               |||      ||||||      :::   :::                  |||   ||||||
Db        1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1195

Qy         392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
                  :::   :::   |||         |||:::      ||||||:::         :::
Db        1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy         412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
                  ||||||      ||||||   |||         :::      |||   |||
Db        1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy         429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438
               ::::::   :::                                 :::   |||:::
Db        1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy         439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
                  |||         ||||||
Db        1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390

Qy         459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
               |||      ||||||||||||:::::::::   |||:::               :::
Db        1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy         479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492
                  |||                         :::::::::|||
Db        1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy         493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
               ||||||:::||||||      |||::::::::::::::::|||   :::|||      |||
Db        1511 CTGAAGAAGCTCTCCTTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy         511 ------------------LysGlnAsnSerThr------------MetPheSerLeuThr 520
                                 ||||||                       |||:::
Db        1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy         521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
                  ::::::            |||   |||   :::::::::|||   :::         |||
Db        1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1690

Qy         541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
               ::::::|||   |||         |||      |||:::               ||||||:::
```

```
Db   1691 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750

Qy    561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db   1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810

Qy    581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db   1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867

Qy    601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
Db   1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927

Qy    620 GlyTyr------------------------------------------TyrLeuGlyMet 625
Db   1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987

Qy    626 PheMetAsnSer------------------------------------------PheAla 631
Db   1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047

Qy    632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
Db   2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107

Qy    652 AspGlyValTyrAspAsnAlaGlu---------------------------------Tyr 660
Db   2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167

Qy    661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db   2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218

Qy    681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
Db   2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269

Qy    698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
Db   2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329

Qy    718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
Db   2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389

Qy    738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
Db   2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTCTTTGTGGAGCTTACAATGCCTTACTCT 2446

Qy    758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
Db   2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506

Qy    778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
Db   2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566

Qy    798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db   2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA--------------------- 2605

Qy    813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db   2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy    833 ThrValThrAlaLeuSer---------------------------------ProThrAla 841
Db   2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy    842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
Db   2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy    857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
Db   2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy    877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
Db   2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy    897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
Db   2885 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy    917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
Db   2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy    937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
Db   3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3064

Qy    957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
Db   3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy    977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
Db   3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy    995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
Db   3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244

Qy   1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
Db   3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy   1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db   3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy   1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
Db   3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy   1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db   3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy   1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
Db   3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy   1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
Db   3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy   1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
Db   3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy   1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
Db   3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy   1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
Db   3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy   1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
Db   3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838
```

```
Qy       1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
              |||      |||  |||||            |||||      ||||||
Db       3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy       1216 SerProLeuAlaValValGln--------------------------------------- 1222
              :::      :::   |||
Db       3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy       1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
                          |||       ::::::        |||   |||
Db       3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy       1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
              |||   ::::::   ||||||:::                                 :::
Db       4016 CAGACCTCCTTGAAATACAATATT--------------------------------CTC 4042

Qy       1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
                 :::|||   |||          :::   |||:::        :::      |||
Db       4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102

Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
              :::         |||          :::       |||       |||
Db       4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
                 |||   ||||||::::::::::|||   ::::::|||||||||:::
Db       4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
                 :::   :::            |||      |||   :::   |||
Db       4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
                 :::   :::|||||||||   |||:::                :::
Db       4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
              ::::::::      |||   :::   :::   |||   |||   :::   |||||||||
Db       4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
              |||         |||:::      ||||||:::
Db       4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438
```

RESULT 8
US-10-331-496A-19
; Sequence 19, Application US/10331496A
; Publication No. US20030228305A1
; GENERAL INFORMATION:
; APPLICANT: FRANTZ,GRETCHEN
; APPLICANT: HILLAN,KENNETH J.
; APPLICANT: PHILLIPS,HEIDI S.
; APPLICANT: POLAKIS,PAUL
; APPLICANT: SMITH,VICTORIA
; APPLICANT: SPENCER,SUSAN D.
; APPLICANT: WILLIAMS,P. MICKEY
; APPLICANT: WU,THOMAS D.
; APPLICANT: ZHANG,ZEMIN
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P5014R1-PCT
; CURRENT APPLICATION NUMBER: US/10/331,496A
; CURRENT FILING DATE: 2002-12-30
; PRIOR APPLICATION NUMBER: US 60/345,444
; PRIOR FILING DATE: 2002-01-02
; PRIOR APPLICATION NUMBER: US 60/351,885
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: US 60/360,066
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: US 60/362,004
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/366,869
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: US 60/366,284
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: US 60/368,679
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/404,809
; PRIOR FILING DATE: 2002-08-19
; PRIOR APPLICATION NUMBER: US 60/405,645
; PRIOR FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 95
; SEQ ID NO 19
; LENGTH: 4577
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-331-496A-19

Alignment Scores:

Pred. No.:	3.56e-133	Length:	4577
Score:	1440.00	Matches:	451
Percent Similarity:	46.20%	Conservative:	266
Best Local Similarity:	29.06%	Mismatches:	551
Query Match:	19.60%	Indels:	284
DB:	12	Gaps:	59

US-10-020-095-4 (1-1428) x US-10-331-496A-19 (1-4577)

```
Qy         12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
              ||||||   :::                  :::|||:::      |||:::::::::|||
Db         86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy         32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
                 |||   ::::::                        |||      ||||||   :::
Db        146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy         52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
                       |||||||||   |||   |||                  |||   ::::::
Db        203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy         72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
                 :::   ||||||||||||      |||               :::|||
Db        263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy         90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
                    ::::::|||   :::|||::::::|||                 ::::::|||
Db        311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
                 |||   ||||||:::            |||         |||   :::      :::
Db        365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
              :::         ||||||:::||||||||||||::::::|||||||||   |||   |||
Db        416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy        147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
              |||||||||:::|||::::::      :::|||   |||            :::   :::
Db        476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy        167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
              :::   |||:::|||||||||   |||   |||   ||||||   |||   |||   :::
Db        536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy        186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
                 |||:::         :::|||   |||||||||   |||         |||   :::
Db        596 GGTGGCCTCAAGCAATTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy        206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
              :::   ||||||               :::|||   :::      |||   |||   |||
Db        656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy        224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
              :::||||||||||||||||||||||   :::         |||         ::::::
Db        713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772
```

```
Qy        244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
              ::: :::|||  :::::           |||||||||||||||||||||     |||
Db        773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGGAAGCCTGTCCCTGGACAT 832

Qy        264 ValThrLeuThrPheLeuProLeu------------------------------------- 271
              ||||||::::::
Db        833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy        272 ---------------------------------------------SerPheTrpGlyLys 276
                                                          |||:::   :::
Db        893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952

Qy        277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
                 |||      ||||||   |||::::::                           :::
Db        953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 988

Qy        297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
                 |||||||||                           |||             :::
Db        989 TATGAAATGAAA---------------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy        317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
                    |||   |||   ||||||                    :::|||   ::::::
Db       1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy        332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
              |||   :::|||      |||      :::      |||:::|||              |||
Db       1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy        352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
                 ||||||                                                :::
Db       1133 CCCTTCTTTGGG---------------------------------------------CAG 1147

Qy        372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
              |||      ||||||      :::   :::                |||   ||||||
Db       1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1195

Qy        392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
              :::   :::      |||      |||:::      ||||||:::                :::
Db       1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy        412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
                       ||||||      ||||||   |||      :::      |||   |||
Db       1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy        429 LysIleGluPhePro------------------------------IleLeuGluAspSer 438
              ::::::   :::                                    :::   |||:::
Db       1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy        439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
                 |||               ||||||
Db       1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390

Qy        459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
              |||      ||||||||||||:::::::::      |||:::                :::
Db       1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy        479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492
                 |||                                 ::::::::|||
Db       1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy        493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
              ||||||:::||||||      |||:::::::::::::::::::|||   :::|||      |||
Db       1511 CTGAAGAAGCTCTCCTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy        511 ------------------LysGlnAsnSerThr------------MetPheSerLeuThr 520
                                |||||                          |||:::
Db       1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
                 ::::::            |||   |||   :::::::::|||      :::      |||
Db       1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTACCTACCGGG 1690

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ::::::|||   |||      |||      |||:::                ||||||:::
Db       1691 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              |||   :::|||      ::::::   |||:::                ||||||::::::
Db       1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
                 |||   |||:::      :::   |||||||||:::||||||   ||||||
Db       1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
                 :::::::::      :::|||   :::   |||   |||                |||
Db       1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927

Qy        620 GlyTyr---------------------------------------TyrLeuGlyMet 625
              |||:::                                                 :::
Db       1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987

Qy        626 PheMetAsnSer-------------------------------------------PheAla 631
              ::::::|||                                           ::::::
Db       1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047

Qy        632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
                 ::::::   ||||||         |||::::::   :::   |||         :::
Db       2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107

Qy        652 AspGlyValTyrAspAsnAlaGlu-------------------------------Tyr 660
                             |||                                      |||
Db       2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
                       |||                          :::|||   :::   :::
Db       2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218

Qy        681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
                       ||||||         |||||||||:::||||||||||||||||||||||
Db       2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269

Qy        698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
                 :::                  |||      |||||||||||||||:::|||
Db       2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329

Qy        718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
              |||   |||   |||      :::|||      |||   |||||||||:::::::::
Db       2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389

Qy        738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
              |||      |||:::|||||||||||||||||||||:::      |||      :::|||||||||
Db       2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446

Qy        758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
              |||||||||||||||      |||   |||:::   |||:::   |||||||||
Db       2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506

Qy        778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
                 |||   |||   :::|||   |||         |||         :::
Db       2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566

Qy        798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
              |||               |||   |||   |||   |||
Db       2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA--------------------- 2605

Qy        813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
                          ||||||   :::   :::   |||      ||||||   :::
Db       2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy        833 ThrValThrAlaLeuSer--------------------------------ProThrAla 841
              ||||||:::|||   :::                                |||:::
Db       2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy        842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
                             |||   |||   :::   :::||||||:::   ||||||:::
```

```
Db       2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy        857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
              ||||||      :::   :::  ||||||        :::         :::
Db       2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy        877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
                 :::  ||||||           |||||||||  |||        |||    |||
Db       2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy        897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
              ::::::   :::||||||:::||||||    ::::::          :::|||::::::
Db       2885 GTCTCAGTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy        917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
              |||||||||||||||||||||||||||||||:::   |||||||||||||||||:::|||
Db       2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy        937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
              |||||||||    :::   :::|||||||||       :::|||   ||||||:::   :::
Db       3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3064

Qy        957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
              :::         ||||||||||||:::|||   |||:::      |||||||||:::|||
Db       3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy        977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
              ||||||      |||          :::   |||:::||||||||||:::||||||||||:::
Db       3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy        995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
                 |||   :::|||     ||||||   ||||||:::    :::         :::
Db       3185 ACTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244

Qy       1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
                 ||||||         ||||||   ||||||   |||         |||   ::::::
Db       3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy       1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
              ::::::   ::::::||||||   :::         |||||||||:::|||||||||
Db       3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy       1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
              :::||||||                     :::      |||:::   ::::::
Db       3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy       1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
              |||||||||                  |||      |||   |||   ||||||
Db       3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy       1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
                 ||||||:::   ||||||   :::      |||:::      |||   ||||||
Db       3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy       1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
              |||      |||         |||   ::::::            |||
Db       3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy       1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
                    ::::::   :::         ::::::|||:::       |||   ::::::
Db       3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy       1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
              |||:::   :::|||   ||||||:::      :::|||         |||     ||||||
Db       3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy       1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
              |||   :::               |||::::::|||:::::::::|||:::|||:::
Db       3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy       1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
              ||||||||:::||||||||||||||||   |||||||||   |||||||||::::::
Db       3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy       1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
              |||         |||   ||||||                  ||||||   ||||||
Db       3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy       1216 SerProLeuAlaValValGln--------------------------------------- 1222
              :::      :::      |||
Db       3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy       1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
                          |||      ::::::            |||   |||
Db       3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy       1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
              |||   ::::::   ||||||:::                                   :::
Db       4016 CAGACCTCCTTGAAATACAATATT--------------------------------CTC 4042

Qy       1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
                 :::|||   |||      :::   |||:::          :::      |||
Db       4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102

Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
              :::         |||         :::      |||   |||      |||
Db       4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
                 |||   ||||||:::::::::|||   ::::::|||||||||:::
Db       4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
                 :::   :::              |||      |||   :::   |||
Db       4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
                 :::   :::||||||||| |||:::                  :::
Db       4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
              :::::::::      |||   :::   :::   |||   |||   :::   |||||||||
Db       4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
              |||         |||:::      ||||||:::
Db       4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438
```

```
RESULT 9
US-09-873-319-408
; Sequence 408, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  APPLICANT: Kulkarni, Prakash
;  APPLICANT: Getzenberg, Robert H.
;  APPLICANT: Waga, Iwao
;  APPLICANT: Yamamoto, Jun
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
;  TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-US
;  CURRENT APPLICATION NUMBER: US/09/873,319A
;  CURRENT FILING DATE: 2001-06-05
;  EARLIER APPLICATION NUMBER: US 60/223,323
;  EARLIER FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 755
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 408
;   LENGTH: 4577
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134324A1 M11313
US-09-873-319-408

Alignment Scores:
```

```
Pred. No.:                  3.56e-133     Length:          4577
Score:                      1440.00       Matches:         451
Percent Similarity:         46.20%        Conservative:    266
Best Local Similarity:      29.06%        Mismatches:      551
Query Match:                19.60%        Indels:          284
DB:                         13            Gaps:            59

US-10-020-095-4 (1-1428) x US-09-873-319-408 (1-4577)

Qy       12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db       86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy       32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
Db      146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy       52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
Db      203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy       72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
Db      263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy       90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
Db      311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy      107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db      365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy      127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
Db      416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy      147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db      476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy      167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
Db      536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy      186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db      596 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy      206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db      656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy      224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db      713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772

Qy      244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db      773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGAAGCCTGTCCCTGGACAT 832

Qy      264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
Db      833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy      272 ---------------------------------------------SerPheTrpGlyLys 276
Db      893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952

Qy      277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db      953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 988

Qy      297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
Db      989 TATGAAATGAAA------------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy      317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
Db     1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy      332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db     1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy      352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
Db     1133 CCCTTCTTTGGG---------------------------------------------CAG 1147

Qy      372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
Db     1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1195

Qy      392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
Db     1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy      412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
Db     1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy      429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438
Db     1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy      439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db     1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390

Qy      459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db     1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy      479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492
Db     1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy      493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
Db     1511 CTGAAGAAGCTCTCCTTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy      511 ------------------LysGlnAsnSerThr------------MetPheSerLeuThr 520
Db     1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy      521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db     1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1690

Qy      541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db     1691 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750

Qy      561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db     1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810

Qy      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db     1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867

Qy      601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
Db     1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927

Qy      620 GlyTyr------------------------------------------TyrLeuGlyMet 625
```

```
Db       1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987

Qy        626 PheMetAsnSer-----------------------------------------PheAla 631

Db       1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047

Qy        632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651

Db       2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107

Qy        652 AspGlyValTyrAspAsnAlaGlu---------------------------------Tyr 660

Db       2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680

Db       2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218

Qy        681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697

Db       2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269

Qy        698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717

Db       2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329

Qy        718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737

Db       2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389

Qy        738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757

Db       2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446

Qy        758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777

Db       2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506

Qy        778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797

Db       2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566

Qy        798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812

Db       2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA--------------------- 2605

Qy        813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832

Db       2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy        833 ThrValThrAlaLeuSer---------------------------------ProThrAla 841

Db       2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy        842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856

Db       2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy        857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876

Db       2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy        877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896

Db       2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy        897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916

Db       2885 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy        917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936

Db       2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy        937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956

Db       3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3064

Qy        957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976

Db       3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy        977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994

Db       3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy        995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014

Db       3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244

Qy       1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034

Db       3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy       1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054

Db       3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy       1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074

Db       3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy       1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088

Db       3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy       1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107

Db       3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy       1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124

Db       3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy       1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142

Db       3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy       1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159

Db       3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy       1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175

Db       3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy       1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195

Db       3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy       1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215

Db       3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy       1216 SerProLeuAlaValValGln--------------------------------------- 1222

Db       3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy       1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238

Db       3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy       1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258

Db       4016 CAGACCTCCTTGAAATACAATATT---------------------------------CTC 4042
```

```
Qy       1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
                 :::|||  |||          :::   |||:::      :::      |||
Db       4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102

Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
              :::         |||          :::      |||  |||    |||
Db       4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
                 |||   ||||||:::::::::|||    ::::::|||||||||:::
Db       4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
                 :::   :::               |||      |||    :::   |||
Db       4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
                 :::   :::|||||||||   |||:::                :::
Db       4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
              :::::::::         |||   :::   :::   |||   |||   :::   ||||||||
Db       4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
              |||            |||:::      ||||||:::
Db       4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438
```

RESULT 10
US-09-960-706-654
; Sequence 654, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
; TITLE OF INVENTION: Gene Expression Profiles
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 654
; LENGTH: 4577
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M11313
US-09-960-706-654

Alignment Scores:

Pred. No.:	3.56e-133	Length:	4577
Score:	1440.00	Matches:	451
Percent Similarity:	46.20%	Conservative:	266
Best Local Similarity:	29.06%	Mismatches:	551
Query Match:	19.60%	Indels:	284
DB:	13	Gaps:	59

US-10-020-095-4 (1-1428) x US-09-960-706-654 (1-4577)

```
Qy         12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
              ||||||   :::                  :::|||:::      |||:::::::::|||
Db         86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy         32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
                 |||   ::::::                   |||      ||||||      :::
Db        146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy         52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
                                   |||||||||   |||   |||               |||   ::::::
Db        203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy         72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
                 :::   ||||||||||||         |||            ::::|||
Db        263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy         90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
                    ::::::|||   :::|||::::::|||                 ::::::|||
Db        311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
                 |||   ||||||:::         |||         |||   :::      :::
Db        365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
              :::            ||||||:::|||||||||||||::::::|||||||||   |||   |||
Db        416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy        147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
              |||||||||:::|||::::::         :::|||   |||               :::   :::
Db        476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy        167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
              :::   |||:::|||||||||   |||   |||   ||||||   |||   |||   :::
Db        536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy        186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
                 |||:::            :::|||   |||||||||   |||      |||   :::
Db        596 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy        206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
              :::   ||||||               :::|||   :::      |||   |||   |||
Db        656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy        224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
              :::||||||||||||||||||||||||   :::         |||         ::::::
Db        713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772

Qy        244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
              :::   :::|||   ::::::            ||||||||||||||||||||   |||
Db        773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGGAAGCCTGTCCCTGGACAT 832

Qy        264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
              ||||||::::::
Db        833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy        272 ------------------------------------------SerPheTrpGlyLys 276
                                                           |||:::   :::
Db        893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952

Qy        277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
                 |||         ||||||   |||::::::                        :::
Db        953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 988

Qy        297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
              |||||||||                              |||            :::
Db        989 TATGAAATGAAA------------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy        317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
                 |||   |||   ||||||                     :::|||   ::::::
Db       1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy        332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
              |||   :::|||   |||      :::      |||:::|||            |||
Db       1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy        352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
                 ||||||                                                 :::
Db       1133 CCCTTCTTTGGG---------------------------------------------CAG 1147

Qy        372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
              |||         ||||||      :::   :::               |||   ||||||
```

```
Db   1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA-----------AATAAAGTCATA 1195

Qy    392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
Db   1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy    412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
Db   1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy    429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438
Db   1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy    439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db   1367 GAAGAGGCACATCACACTGCTTAT----------------------------------- 1390

Qy    459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db   1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy    479 ValGlySer-----------ProPheGluLeuValValSerGlyAsnLys------Arg 492
Db   1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy    493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
Db   1511 CTGAAGAAGCTCTCCTTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy    511 ------------------LysGlnAsnSerThr-----------MetPheSerLeuThr 520
Db   1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy    521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db   1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1690

Qy    541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db   1691 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750

Qy    561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db   1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810

Qy    581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db   1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867

Qy    601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
Db   1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927

Qy    620 GlyTyr-----------------------------------------TyrLeuGlyMet 625
Db   1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987

Qy    626 PheMetAsnSer------------------------------------------PheAla 631
Db   1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047

Qy    632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
Db   2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107

Qy    652 AspGlyValTyrAspAsnAlaGlu--------------------------------Tyr 660
Db   2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167

Qy    661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db   2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218

Qy    681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
Db   2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269

Qy    698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
Db   2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329

Qy    718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
Db   2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389

Qy    738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
Db   2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446

Qy    758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
Db   2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506

Qy    778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
Db   2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566

Qy    798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db   2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA-------------------- 2605

Qy    813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db   2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy    833 ThrValThrAlaLeuSer---------------------------------ProThrAla 841
Db   2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy    842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
Db   2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy    857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
Db   2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy    877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
Db   2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy    897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
Db   2885 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy    917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
Db   2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy    937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
Db   3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3064

Qy    957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
Db   3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy    977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
Db   3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy    995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
Db   3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244
```

```
Qy       1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
Db       3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy       1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db       3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy       1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
Db       3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy       1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db       3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy       1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
Db       3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy       1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
Db       3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy       1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
Db       3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy       1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
Db       3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy       1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
Db       3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy       1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
Db       3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy       1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
Db       3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy       1216 SerProLeuAlaValValGln--------------------------------------- 1222
Db       3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy       1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
Db       3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy       1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
Db       4016 CAGACCTCCTTGAAATACAATATT---------------------------------CTC 4042

Qy       1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
Db       4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102

Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
Db       4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
Db       4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
Db       4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
Db       4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
Db       4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
Db       4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438
```

RESULT 11
US-10-292-081A-3
; Sequence 3, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGL
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4577
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-292-081A-3

Alignment Scores:

Pred. No.:	3.56e-133	Length:	4577
Score:	1440.00	Matches:	451
Percent Similarity:	46.20%	Conservative:	266
Best Local Similarity:	29.06%	Mismatches:	551
Query Match:	19.60%	Indels:	284
DB:	13	Gaps:	59

US-10-020-095-4 (1-1428) x US-10-292-081A-3 (1-4577)

```
Qy         12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db         86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy         32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
Db        146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy         52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
Db        203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy         72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
Db        263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy         90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
Db        311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db        365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
Db        416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475
```

```
Qy    147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db    476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy    167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
Db    536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy    186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db    596 GGTGGCCTCAAGCAATTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy    206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db    656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy    224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db    713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772

Qy    244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db    773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGAAGCCTGTCCCTGGACAT 832

Qy    264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
Db    833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy    272 ------------------------------------------------SerPheTrpGlyLys 276
Db    893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952

Qy    277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db    953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG-----------------AAGGAG 988

Qy    297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
Db    989 TATGAAATGAAA-----------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy    317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
Db   1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy    332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db   1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy    352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
Db   1133 CCCTTCTTTGGG---------------------------------------------CAG 1147

Qy    372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
Db   1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1195

Qy    392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
Db   1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy    412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
Db   1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy    429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438
Db   1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy    439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db   1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390

Qy    459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db   1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy    479 ValGlySer-----------ProPheGluLeuValValSerGlyAsnLys------Arg 492
Db   1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy    493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
Db   1511 CTGAAGAAGCTCTCCTTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy    511 -----------------LysGlnAsnSerThr------------MetPheSerLeuThr 520
Db   1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy    521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db   1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1690

Qy    541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db   1691 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750

Qy    561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db   1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810

Qy    581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db   1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867

Qy    601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
Db   1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927

Qy    620 GlyTyr------------------------------------------TyrLeuGlyMet 625
Db   1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987

Qy    626 PheMetAsnSer------------------------------------------PheAla 631
Db   1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047

Qy    632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
Db   2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107

Qy    652 AspGlyValTyrAspAsnAlaGlu---------------------------------Tyr 660
Db   2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167

Qy    661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db   2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218

Qy    681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
Db   2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269

Qy    698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
Db   2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329

Qy    718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
Db   2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389

Qy    738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
Db   2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446

Qy    758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
```

```
Db   2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506

Qy    778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
Db   2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566

Qy    798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db   2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA--------------------- 2605

Qy    813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db   2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy    833 ThrValThrAlaLeuSer--------------------------------ProThrAla 841
Db   2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy    842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
Db   2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy    857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
Db   2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy    877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
Db   2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy    897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
Db   2885 GTCTCAGTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy    917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
Db   2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy    937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
Db   3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3064

Qy    957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
Db   3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy    977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
Db   3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy    995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
Db   3185 ACTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244

Qy   1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
Db   3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy   1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db   3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy   1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
Db   3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy   1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db   3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy   1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
Db   3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy   1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
Db   3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy   1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
Db   3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy   1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
Db   3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy   1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
Db   3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy   1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
Db   3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy   1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
Db   3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy   1216 SerProLeuAlaValValGln--------------------------------------- 1222
Db   3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy   1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
Db   3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy   1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258
Db   4016 CAGACCTCCTTGAAATACAATATT--------------------------------CTC 4042

Qy   1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
Db   4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102

Qy   1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
Db   4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy   1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
Db   4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy   1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
Db   4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy   1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
Db   4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy   1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
Db   4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy   1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
Db   4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438
```

RESULT 12
US-10-292-081A-5
; Sequence 5, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi

```
;  APPLICANT:  Xin Wang
;  APPLICANT:  Randolph E. Tanzi
;  APPLICANT:  Lars Bertram
;  APPLICANT:  Aleister J. Saunders
;  TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLC
;  FILE REFERENCE: 37481-3323
;  CURRENT APPLICATION NUMBER: US/10/292,081A
;  CURRENT FILING DATE:  2002-11-08
;  PRIOR APPLICATION NUMBER: 60/337434
;  PRIOR FILING DATE: 2001-11-09
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 4577
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-292-081A-5

Alignment Scores:
Pred. No.:              3.56e-133     Length:        4577
Score:                  1440.00       Matches:       451
Percent Similarity:     46.20%        Conservative:  266
Best Local Similarity:  29.06%        Mismatches:    551
Query Match:            19.60%        Indels:        284
DB:                     13            Gaps:          59

US-10-020-095-4 (1-1428) x US-10-292-081A-5 (1-4577)

Qy        12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
             ||||||   :::                     :::|||:::       |||:::::::::|||
Db        86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy        32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
                |||    ::::::                     |||       ||||||    :::
Db       146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy        52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
                         |||||||||   |||   |||          |||          ::::::
Db       203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy        72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
                :::   |||||||||||       |||            :::|||
Db       263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy        90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
                   ::::::|||   :::|||::::::|||                 ::::::|||
Db       311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy       107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
                |||   ||||||:::          |||          |||    :::      :::
Db       365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy       127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
             :::            ||||||:::||||||||||||:::::::||||||||   |||    |||
Db       416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy       147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
             |||||||||:::|||::::::          :::|||   |||            :::   :::
Db       476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy       167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
             :::      |||:::||||||||||||   |||   |||   ||||||   |||   |||   :::
Db       536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy       186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
                |||:::             :::|||   |||||||||   |||      |||      :::
Db       596 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy       206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
             :::      ||||||              :::|||   :::      |||   |||   |||
Db       656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy       224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
             :::||||||||||||||||||||   :::      |||            ::::::
Db       713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772

Qy       244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
             :::   :::|||   ::::::          ||||||||||||||||||   |||
Db       773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGAAGCCTGTCCCTGGACAT 832

Qy       264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
             ||||||::::::
Db       833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy       272 ------------------------------------------SerPheTrpGlyLys 276
                                                          |||:::   :::
Db       893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952

Qy       277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
                |||      ||||||   |||::::::                              :::
Db       953 GTAAAA-----ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 988

Qy       297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
                |||||||||                              |||          :::
Db       989 TATGAAATGAAA------------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy       317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
                |||   |||   ||||||                          :::|||   ::::::
Db      1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy       332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
             |||   :::|||    |||      :::      |||:::|||               |||
Db      1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy       352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
                ||||||                                                  :::
Db      1133 CCCTTCTTTGGG---------------------------------------------CAG 1147

Qy       372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
             |||          ||||||      :::   :::                |||   ||||||
Db      1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1195

Qy       392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
                :::   :::   |||      |||:::      ||||||:::          :::
Db      1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy       412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
                   ||||||      ||||||   |||          :::      |||   |||
Db      1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy       429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438
             ::::::   :::                                :::   |||:::
Db      1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy       439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
                |||         ||||||
Db      1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390

Qy       459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
             |||      ||||||||||||:::::::::   |||:::                  :::
Db      1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy       479 ValGlySer-----------ProPheGluLeuValValSerGlyAsnLys------Arg 492
                |||                             ::::::::::|||
Db      1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy       493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
             ||||||:::||||||         |||::::::::::::::|||      :::|||      |||
Db      1511 CTGAAGAAGCTCTCCTTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy       511 -----------------LysGlnAsnSerThr-----------MetPheSerLeuThr 520
                                 ||||||                        |||:::
Db      1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy       521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
                ::::::         |||   |||      ::::::::|||    :::            |||
```

```
Db   1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTACCTACCGGG 1690
Qy    541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db   1691 GACGTGATTGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750
Qy    561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db   1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810
Qy    581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db   1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867
Qy    601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
Db   1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927
Qy    620 GlyTyr------------------------------------------TyrLeuGlyMet 625
Db   1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987
Qy    626 PheMetAsnSer------------------------------------------PheAla 631
Db   1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047
Qy    632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
Db   2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107
Qy    652 AspGlyValTyrAspAsnAlaGlu---------------------------------Tyr 660
Db   2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167
Qy    661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db   2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218
Qy    681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
Db   2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269
Qy    698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
Db   2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329
Qy    718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
Db   2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389
Qy    738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
Db   2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446
Qy    758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
Db   2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506
Qy    778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
Db   2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566
Qy    798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db   2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA--------------------- 2605
Qy    813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db   2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653
Qy    833 ThrValThrAlaLeuSer---------------------------------ProThrAla 841
Db   2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy    842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
Db   2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773
Qy    857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
Db   2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827
Qy    877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
Db   2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884
Qy    897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
Db   2885 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944
Qy    917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
Db   2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004
Qy    937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
Db   3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3064
Qy    957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
Db   3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124
Qy    977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
Db   3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184
Qy    995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
Db   3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244
Qy   1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
Db   3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304
Qy   1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db   3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364
Qy   1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
Db   3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424
Qy   1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db   3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481
Qy   1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
Db   3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541
Qy   1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
Db   3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601
Qy   1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
Db   3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661
Qy   1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
Db   3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718
Qy   1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
Db   3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778
```

```
Qy       1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
              ||||||||:::|||||||||||||||    |||||||||  ||||||||||::::::
Db       3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy       1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
              |||         |||  ||||||                   ||||||    ||||||
Db       3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy       1216 SerProLeuAlaValValGln---------------------------------------- 1222
              :::      :::      |||
Db       3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy       1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
                          |||         ::::::       |||   |||
Db       3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy       1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
              |||   ::::::   ||||||:::                                :::
Db       4016 CAGACCTCCTTGAAATACAATATT--------------------------------CTC 4042

Qy       1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
                 :::|||   |||         :::  |||:::         :::        |||
Db       4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102

Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
              :::         |||         :::      |||   |||      |||
Db       4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
                    |||   ||||||:::::::::|||   ::::::|||||||||:::
Db       4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
                 :::   :::            |||      |||   :::   |||
Db       4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
                 :::   :::||||||||| |||:::                     :::
Db       4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
              :::::::::      |||   :::   :::   |||   |||   :::   ||||||||
Db       4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
              |||            |||:::      ||||||:::
Db       4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438
```

RESULT 13
US-10-076-816-56
; Sequence 56, Application US/10076816
; Publication No. US20030056244A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodrguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT APPLICATION NUMBER: US/10/076,816
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/269,188
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 4577
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-076-816-56

```
Alignment Scores:
Pred. No.:               3.56e-133     Length:         4577
Score:                   1440.00       Matches:        451
Percent Similarity:      46.20%        Conservative:   266
Best Local Similarity:   29.06%        Mismatches:     551
Query Match:             19.60%        Indels:         284
DB:                      15            Gaps:           59

US-10-020-095-4 (1-1428) x US-10-076-816-56 (1-4577)

Qy         12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
              ||||||   :::                  :::|||:::       |||:::::::::|||
Db         86 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 145

Qy         32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
                 |||   ::::::                       |||      ||||||   :::
Db        146 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 202

Qy         52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
                       |||||||||   |||   |||                  |||   ::::::
Db        203 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 262

Qy         72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
                 :::   ||||||||||||      |||         :::|||
Db        263 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 310

Qy         90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
                 ::::::|||   :::|||::::::|||                  ::::::|||
Db        311 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 364

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
                 |||   ||||||:::               |||      |||   :::   :::
Db        365 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 415

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
              :::            ||||||:::||||||||||||::::::|||||||||   |||   |||
Db        416 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 475

Qy        147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
              |||||||||:::|||::::::         :::|||   |||            :::   :::
Db        476 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 535

Qy        167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
              :::      |||:::|||||||||   |||   |||   ||||||   |||   |||   :::
Db        536 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 595

Qy        186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
                 |||:::            :::|||   |||||||||   |||         |||   :::
Db        596 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 655

Qy        206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
              :::   ||||||               :::|||   :::         |||   |||   |||
Db        656 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 712

Qy        224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
              :::|||||||||||||||||||||||||   :::         |||            ::::::
Db        713 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 772

Qy        244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
              :::   :::|||   ::::::               ||||||||||||||||||||||   |||
Db        773 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGGAAGCCTGTCCCTGGACAT 832

Qy        264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
              ||||||::::::
Db        833 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 892

Qy        272 ---------------------------------------------SerPheTrpGlyLys 276
                                                           |||:::   :::
Db        893 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 952
```

```
Qy        277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
                  |||       ||||||  |||::::::                              :::
Db        953 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 988

Qy        297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
                 |||||||||                          |||              :::
Db        989 TATGAAATGAAA------------------------CTTCACACTGAGGCCCAGATCCAA 1024

Qy        317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
                    |||   |||    ||||||                 :::|||    ::::::
Db       1025 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1084

Qy        332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
              |||    :::|||      |||      :::      |||:::|||              |||
Db       1085 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1132

Qy        352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
              ||||||                                                      :::
Db       1133 CCCTTCTTTGGG---------------------------------------------CAG 1147

Qy        372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
              |||       ||||||      :::   :::                |||   ||||||
Db       1148 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1195

Qy        392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
                 :::   :::   |||         |||:::      ||||||:::            :::
Db       1196 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1246

Qy        412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
                       ||||||      ||||||   |||       :::       |||   |||
Db       1247 CATGGCCTTGTACAGTTCTCTATCAACACCACCAACGTTATGGGTACCTCTCTTACTGTT 1306

Qy        429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438
              ::::::   :::                                 :::     |||:::
Db       1307 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1366

Qy        439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
                 |||         ||||||
Db       1367 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1390

Qy        459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
              |||         ||||||||||||:::::::::      |||:::
Db       1391 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1450

Qy        479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492
                 |||                        :::::::::|||
Db       1451 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1510

Qy        493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
              ||||||:::||||||      |||::::::::::::::::|||    :::|||       |||
Db       1511 CTGAAGAAGCTCTCCTTTTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1570

Qy        511 ------------------LysGlnAsnSerThr------------MetPheSerLeuThr 520
                                ||||||                          |||:::
Db       1571 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1630

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
                 ::::::          |||   |||    ::::::::|||      :::          |||
Db       1631 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1690

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
              ::::::|||   |||         |||      |||:::                ||||||:::
Db       1691 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1750

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
              |||    :::|||      ::::::      |||:::              ||||||::::::
Db       1751 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1810

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
                 |||   |||:::          :::   |||||||||:::||||||      ||||||
Db       1811 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1867

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
                    :::::::::              :::|||   :::      |||   |||           |||
Db       1868 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1927

Qy        620 GlyTyr---------------------------------------TyrLeuGlyMet 625
              |||:::                                                :::
Db       1928 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 1987

Qy        626 PheMetAsnSer-----------------------------------------PheAla 631
              ::::::|||                                               ::::::
Db       1988 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2047

Qy        632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
                 ::::::      ||||||         |||::::::   :::   |||      :::
Db       2048 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2107

Qy        652 AspGlyValTyrAspAsnAlaGlu-------------------------------Tyr 660
                                   |||                                 |||
Db       2108 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2167

Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
                       |||                     :::|||   :::   :::
Db       2168 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2218

Qy        681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
                    ||||||         |||||||||:::||||||||||||||||||||
Db       2219 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2269

Qy        698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
                 :::                     |||      ||||||||||||||||:::|||
Db       2270 TTGGTGGTGGTAAACTCAGCAGGGGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2329

Qy        718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
              |||   |||   |||      |||    :::||||||||   ||||||||:::::::::
Db       2330 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2389

Qy        738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
              |||         |||:::||||||||||||||||||:::    |||    :::||||||||
Db       2390 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTTACAATGCCTTACTCT 2446

Qy        758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
              |||||||||||||||      |||   |||:::   |||:::   ||||||||
Db       2447 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2506

Qy        778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
                 |||   |||   :::|||   |||      |||         :::
Db       2507 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTTGCTGTCCCAGTGGAGAAG 2566

Qy        798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
              |||               |||   |||   |||   |||
Db       2567 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA--------------------- 2605

Qy        813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
                             ||||||   :::   :::   |||      ||||||   :::
Db       2606 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2653

Qy        833 ThrValThrAlaLeuSer---------------------------------ProThrAla 841
              ||||||:::|||    :::                                   |||:::
Db       2654 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2713

Qy        842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
                                |||   |||      :::   :::||||||:::   ||||||:::
Db       2714 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2773

Qy        857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
              ||||||      :::   :::   ||||||         :::            :::
Db       2774 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2827

Qy        877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
                 :::   ||||||         |||||||||   |||         |||   |||
Db       2828 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2884

Qy        897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
              ::::::   :::||||||:::||||||         ::::::            :::|||::::::
```

```
Db       2885 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2944

Qy        917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
              |||||||||||||||||||||||||||||:::  ·  |||||||||||||||||:::|||
Db       2945 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3004

Qy        937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
              |||||||||   :::    :::|||||||||        :::|||   ||||||:::    :::
Db       3005 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3064

Qy        957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
              :::         |||||||||||:::|||   |||:::          ||||||||:::|||
Db       3065 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3124

Qy        977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
              ||||||      |||       :::     |||:::||||||||:::|||||||||||:::
Db       3125 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3184

Qy        995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
                 |||   :::|||      ||||||   ||||||:::        :::   :::
Db       3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3244

Qy       1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
                 ||||||          ||||||   ||||||   |||         |||   ::::::
Db       3245 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3304

Qy       1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
              ::::::   ::::::|||||     :::        |||||||||:::||||||||
Db       3305 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3364

Qy       1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
              :::||||||                     :::        |||:::    ::::::
Db       3365 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3424

Qy       1075 LeuGluSer------------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
              |||||||||                  |||      |||   |||      ||||||
Db       3425 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3481

Qy       1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
                 ||||||:::   ||||||   :::      |||:::      |||   ||||||
Db       3482 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3541

Qy       1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
              |||      |||         |||   ::::::            |||
Db       3542 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3601

Qy       1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
                 ::::::   :::         ::::::|||:::       |||   ::::::
Db       3602 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3661

Qy       1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
              |||:::   :::|||   ||||||:::   :::|||      |||         ||||||
Db       3662 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3718

Qy       1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
              |||   :::               |||::::::|||:::::::::|||:::|||:::
Db       3719 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3778

Qy       1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
              |||||||||:::|||||||||||||||      |||||||||   |||||||||::::::
Db       3779 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3838

Qy       1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
              |||         |||   ||||||                ||||||      ||||||
Db       3839 GCCGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3895

Qy       1216 SerProLeuAlaValValGln--------------------------------------- 1222
              :::      :::   |||
Db       3896 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAATCGCCTGTTACTGCAGCAGGTCTCA 3955

Qy       1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
                          |||         ::::::         |||   |||
Db       3956 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4015

Qy       1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
              |||   ::::::   ||||||:::                                 :::
Db       4016 CAGACCTCCTTGAAATACAATATT--------------------------------CTC 4042

Qy       1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
                 :::|||   |||         :::   |||:::         :::      |||
Db       4043 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4102

Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
              :::         |||         :::      |||   |||         |||
Db       4103 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4162

Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
                 |||   ||||||:::::::::|||   ::::::|||||||||:::
Db       4163 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4222

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
                 :::   :::               |||      |||   :::   |||
Db       4223 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4282

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
                 :::   :::|||||||||   |||:::                  :::
Db       4283 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4342

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
              ::::::::      |||   :::   :::   |||   |||   :::   |||||||||
Db       4343 CTGCAAGATGTCCCAGTAAGAGATCTCAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4402

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
              |||         |||:::      ||||||:::
Db       4403 GAGACGGATGAGTTTGCAATCGCTGAGTACAATGCT 4438
```

```
RESULT 14
US-10-240-965-178
; Sequence 178, Application US/10240965
; Publication No. US20030165924A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  SHIFFMAN, Dov
;  APPLICANT:  SOMOGYI, Roland
;  APPLICANT:  LAWN, Richard M.
;  APPLICANT:  SEILHAMER, Jeffrey J.
;  APPLICANT:  PORTER, Gordon J.
;  APPLICANT:  MIKITA, Thomas
;  APPLICANT:  TAI, Julie
;  TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
;  FILE REFERENCE: PA-0025 PCT
;  CURRENT APPLICATION NUMBER: US/10/240,965
;  CURRENT FILING DATE:  2002-10-04
;  PRIOR APPLICATION NUMBER: 60/195,106
;  PRIOR FILING DATE: 2000-04-05
;  NUMBER OF SEQ ID NOS: 276
;  SOFTWARE: PERL Program
; SEQ ID NO 178
;   LENGTH: 4809
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030165924A1 1000033.6
US-10-240-965-178

Alignment Scores:
Pred. No.:              3.87e-133    Length:        4809
Score:                  1440.00      Matches:       451
Percent Similarity:     46.20%       Conservative:  266
Best Local Similarity:  29.06%       Mismatches:    551
Query Match:            19.60%       Indels:        284
DB:                     13           Gaps:          59

US-10-020-095-4 (1-1428) x US-10-240-965-178 (1-4809)
```

```
Qy         12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db        116 CTCTTGGTCCTCCTGCCCACAGACGCCTCAGTCTCTGGAAAACCGCAGTATATGGTTCTG 175

Qy         32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
Db        176 GTCCCCTCCCTGCTCCACACT---GAGACCACTGAGAAGGGCTGTGTCCTTCTGAGCTAC 232

Qy         52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
Db        233 CTGAATGAGACAGTGACTGTAAGTGCTTCCTTGGAGTCTGTCAGGGGAAACAGGAGCCTC 292

Qy         72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
Db        293 TTCACTGACCTGGAGGCGGAGAATGACGTACTCCACTGTGTCGCCTTC------------ 340

Qy         90 LeuProSerLeuProLeuAsnSerAlaAspGlu---------IleTyrGluLeuArgVal 106
Db        341 ------GCTGTCCCAAAGTCTTCATCCAATGAGGAGGTAATGTTCCTCACTGTCCAAGTG 394

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db        395 AAAGGACCAACCCAAGAA---------TTTAAGAAGCGGACCACAGTGATGGTTAAGAAC 445

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
Db        446 GAGGACAGTCTGGTCTTTGTCCAGACAGACAAATCAATCTACAAACCAGGGCAGACAGTG 505

Qy        147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db        506 AAATTTCGTGTTGTCTCCATGGATGAAAACTTTCACCCCCTGAATGAGTTGATTCCACTA 565

Qy        167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
Db        566 GTATACATTCAGGATCCCAAAGGAAATCGCATCGCACAATGGCAGAGTTTCCAGTTAGAG 625

Qy        186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db        626 GGTGGCCTCAAGCAATTTTCTTTTCCCCTCTCATCAGAGCCCTTCCAGGGCTCCTACAAG 685

Qy        206 IleGlnValGlnValAsnAsp------GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db        686 GTGGTGGTACAGAAGAAATCAGGTGGAAGGACAGAGCAC---CCTTTCACCGTGGAGGAA 742

Qy        224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db        743 TTTGTTCTTCCCAAGTTTGAAGTACAAGTAACAGTGCCAAAGATAATCACCATCTTGGAA 802

Qy        244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db        803 GAAGAGATGAATGTATCAGTGTGTGGCCTATACACATATGGAAGCCTGTCCCTGGACAT 862

Qy        264 ValThrLeuThrPheLeuProLeu------------------------------------ 271
Db        863 GTGACTGTGAGCATTTGCAGAAAGTATAGTGACGCTTCCGACTGCCACGGTGAAGATTCA 922

Qy        272 ------------------------------------------SerPheTrpGlyLys 276
Db        923 CAGGCTTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCCATGGCTGCTTCTATCAGCAA 982

Qy        277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db        983 GTAAAA------ACCAAGGTCTTCCAGCTGAAGAGG------------------AAGGAG 1018

Qy        297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316
Db       1019 TATGAAATGAAA------------------------CTTCACACTGAGGCCCAGATCCAA 1054

Qy        317 SerProGlyProValGluIleLeuThr---------------ThrValThrGluSerVal 331
Db       1055 GAAGAAGGAACAGTGGTGGAATTGACTGGAAGGCAGTCCAGTGAAATCACAAGAACCATA 1114

Qy        332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db       1115 ACCAAACTCTCA---TTTGTGAAAGTGGACTCACACTTTCGACAGGGA---------ATT 1162

Qy        352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
Db       1163 CCCTTCTTTGGG---------------------------------------------CAG 1177

Qy        372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391
Db       1178 GTGCGCCTAGTAGATGGGAAAGGCGTCCCTATACCA------------AATAAAGTCATA 1225

Qy        392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
Db       1226 TTCATC---AGAGGAAATGAAGCAAACTATTAC------TCCAATGCTACCACGGATGAG 1276

Qy        412 MetGluAlaValGln---LysIleAsnTyrThr------ValProGlnSerGlyThrPhe 428
Db       1277 CATGGCCTTGTACAGTTCTCTATCAACACCACCAATGTTATGGGTACCTCTCTTACTGTT 1336

Qy        429 LysIleGluPhePro-----------------------------IleLeuGluAspSer 438
Db       1337 AGGGTCAATTACAAGGATCGTAGTCCCTGTTACGGCTACCAGTGGGTGTCAGAAGAACAC 1396

Qy        439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db       1397 GAAGAGGCACATCACACTGCTTAT------------------------------------ 1420

Qy        459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db       1421 CTTGTGTTCTCCCCAAGCAAGAGCTTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1480

Qy        479 ValGlySer------------ProPheGluLeuValValSerGlyAsnLys------Arg 492
Db       1481 TGTGGCCATACTCAGACAGTCCAGGCACATTATATTCTGAATGGAGGCACCCTGCTGGGG 1540

Qy        493 LeuLysGluLeuSer------TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
Db       1541 CTGAAGAAGCTCTCCTTCTATTATCTGATAATGGCAAAGGGAGGCATTGTCCGAACTGGG 1600

Qy        511 ------------------LysGlnAsnSerThr------------MetPheSerLeuThr 520
Db       1601 ACTCATGGACTGCTTGTGAAGCAGGAAGACATGAAGGGCCATTTTTCCATCTCAATCCCT 1660

Qy        521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db       1661 GTGAAGTCAGACATTGCTCCTGTCGCTCGGTTGCTCATCTATGCTGTTTTACCTACCGGG 1720

Qy        541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db       1721 GACGTGATTGGGGATTCTGCAAAATATGATGTTGAAAATTGTCTGGCCAACAAGGTGGAT 1780

Qy        561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db       1781 TTGAGCTTCAGCCCATCACAAAGTCTCCCAGCCTCACACGCCCACCTGCGAGTCACAGCG 1840

Qy        581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db       1841 GCT---CCTCAGTCCGTCTGCGCCCTCCGTGCTGTGGACCAAAGCGTGCTGCTCATGAAG 1897

Qy        601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619
Db       1898 CCTGATGCTGAGCTCTCGGCGTCCTCGGTTTACAACCTGCTACCAGAAAAGGACCTCACT 1957

Qy        620 GlyTyr------------------------------------------TyrLeuGlyMet 625
Db       1958 GGCTTCCCTGGGCCTTTGAATGACCAGGACGATGAAGACTGCATCAATCGTCATAATGTC 2017

Qy        626 PheMetAsnSer------------------------------------------PheAla 631
Db       2018 TATATTAATGGAATCACATATACTCCAGTATCAAGTACAAATGAAAAGGATATGTACAGC 2077

Qy        632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651
```

```
Db       2078 TTCCTAGAGGACATGGGCTTAAAGGCATTCACCAACTCAAAGATTCGTAAACCCAAAATG 2137
Qy        652 AspGlyValTyrAspAsnAlaGlu--------------------------------Tyr 660
Db       2138 TGTCCACAGCTTCAACAGTATGAAATGCATGGACCTGAAGGTCTACGTGTAGGTTTTTAT 2197
Qy        661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680
Db       2198 GAGTCAGATGTAATGGGAAGAGGCCATGCACGCCTGGTGCATGTTGAAGAG--------- 2248
Qy        681 GlySerSerProHis---------ValArgLysHisPheProGluThrTrpIleTrpLeu 697
Db       2249 ---------CCTCACACGGAGACCGTACGAAAGTACTTCCCTGAGACATGGATCTGGGAT 2299
Qy        698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717
Db       2300 TTGGTGGTGGTAAACTCAGCAGGTGTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATC 2359
Qy        718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737
Db       2360 ACCGAGTGGAAGGCAGGGGCCTTCTGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCC 2419
Qy        738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757
Db       2420 ACT---GCCTCTCTCCGAGCCTTCCAGCCCTTCTTTGTGGAGCTCACAATGCCTTACTCT 2476
Qy        758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777
Db       2477 GTGATTCGTGGAGAGGCCTTCACACTCAAGGCCACGGTCCTAAACTACCTTCCCAAATGC 2536
Qy        778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
Db       2537 ATCCGGGTCAGTGTGCAGCTGGAAGCCTCTCCCGCCTTCCTAGCTGTCCCAGTGGAGAAG 2596
Qy        798 Glu---------------IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db       2597 GAACAAGCGCCTCACTGCATCTGTGCAAACGGGCGGCAA--------------------- 2635
Qy        813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db       2636 ------------ACTGTGTCCTGGGCAGTAACCCCAAAGTCATTAGGAAATGTGAATTTC 2683
Qy        833 ThrValThrAlaLeuSer--------------------------------ProThrAla 841
Db       2684 ACTGTGAGCGCAGAGGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGTGCCTTCAGTT 2743
Qy        842 ---------------SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856
Db       2744 CCTGAACACGGAAGGAAAGACACAGTCATCAAGCCTCTGTTGGTTGAACCTGAAGGACTA 2803
Qy        857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876
Db       2804 GAGAAGGAAACAACATTCAACTCCCTACTTTGTCCATCAGGTGGTGAGGTTTCT------ 2857
Qy        877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896
Db       2858 ---GAAGAATTATCCCTGAAACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCT 2914
Qy        897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916
Db       2915 GTCTCAGTTTTGGGAGACATATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAG 2974
Qy        917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936
Db       2975 ATGCCCTATGGCTGTGGAGAGCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTG 3034
Qy        937 AspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956
Db       3035 GATTATCTAAATGAAACACAGCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTAT 3094
Qy        957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976
Db       3095 CTCAACACTGGTTACCAGAGACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACC 3154

Qy        977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
Db       3155 TTTGGGGAGCGATATGGCAGGAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAG 3214
Qy        995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
Db       3215 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTC 3274
Qy       1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
Db       3275 ATATGGCTCTCCCAGAGGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTC 3334
Qy       1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db       3335 AACAATGCCATAAAGGGAGGAGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATC 3394
Qy       1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
Db       3395 GCCCTTCTGGAGATTCCTCTCACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGC 3454
Qy       1075 LeuGluSer-----------------GluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db       3455 CTGGAGTCAGCCTGGAAGACAGCACAAGAAGGGGACCATGGC---AGCCATGTATATACC 3511
Qy       1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107
Db       3512 AAAGCACTGCTGGCCTATGCTTTTGCCCTGGCAGGTAACCAGGACAAGAGGAAGGAAGTA 3571
Qy       1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp--------- 1124
Db       3572 CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACTCTGTCCATTGGGAGCGCCCT 3631
Qy       1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg------SerLeuAspIle 1142
Db       3632 CAGAAACCCAAGGCACCAGTGGGGCATTTTTACGAACCCCAGGCTCCCTCTGCTGAGGTG 3691
Qy       1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln---------ThrSer 1159
Db       3692 GAGATGACATCCTATGTGCTCCTCGCT---TATCTCACGGCCCAGCCAGCCCCAACCTCG 3748
Qy       1160 GluGlyIlePro------------IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175
Db       3749 GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAG 3808
Qy       1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195
Db       3809 GGCGGTTTCTCCTCCACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGA 3868
Qy       1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
Db       3869 GCAGCCACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACTATCCAGTCTTCAGGG 3925
Qy       1216 SerProLeuAlaValValGln-------------------------------------- 1222
Db       3926 ACATTTTCCAGCAAATTCCAAGTGGACAACAACAACCGCCTGTTACTGCAGCAGGTCTCA 3985
Qy       1223 ------------ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
Db       3986 TTGCCAGAGCTGCCTGGGGAATACAGCATGAAAGTGACAGGAGAAGGATGTGTCTACCTC 4045
Qy       1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIle 1258
Db       4046 CAGACATCCTTGAAATACAATATT--------------------------------CTC 4072
Qy       1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------GluAsnLysAsp 1275
Db       4073 CCAGAAAAGGAAGAGTTCCCCTTTGCTTTAGGAGTGCAGACTCTGCCTCAAACTTGTGAT 4132
Qy       1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly------ 1292
Db       4133 GAACCCAAAGCCCACACCAGCTTCCAAATCTCCCTAAGTGTCAGTTACACAGGGAGCCGC 4192
```

```
Qy       1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
                     |||   ||||||:::::::::|||    ::::::|||||||||:::
Db       4193 TCTGCCTCCAACATGGCGATCGTTGATGTGAAGATGGTCTCTGGCTTCATTCCCCTGAAG 4252

Qy       1312 GluAlaIleSerLeu---------SerGluThrValLysLysValGluTyrAspHisGly 1328
                 :::   :::             |||     |||   :::   |||
Db       4253 CCAACAGTGAAAATGCTTGAAAGATCTAACCATGTGAGCCGGACAGAAGTCAGCAGCAAC 4312

Qy       1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
                 :::   :::|||||||||   |||:::                :::
Db       4313 CATGTCTTGATTTACCTTGATAAGGTGTCAAATCAGACACTGAGCTTGTTCTTCACGGTT 4372

Qy       1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
              :::::::::         |||   :::   :::   |||   |||   :::   |||||||||
Db       4373 CTGCAAGATGTCCCAGTAAGAGATCTGAAACCAGCCATAGTGAAAGTCTATGATTACTAC 4432

Qy       1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
              |||         |||:::      ||||||:::
Db       4433 GAGACGGATGAGTTTGCAATTGCTGAGTACAATGCT 4468
```

RESULT 15
US-09-756-247-3
; Sequence 3, Application US/09756247
; Publication No. US20030180722A1
; GENERAL INFORMATION:
; APPLICANT: Godbole, Shubhada D
; APPLICANT: Boyle, Bryan J
; APPLICANT: Mize, Nancy K
; APPLICANT: Deng, Cenhua
; APPLICANT: Goodrich, Ryle
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Zhou, Ping
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Yeung, George
; APPLICANT: Drmanac, Radoje T .
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE PO
; TITLE OF INVENTION: AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-31CIP
; CURRENT APPLICATION NUMBER: US/09/756,247
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/684,711
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 4527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4527)
US-09-756-247-3

```
Alignment Scores:
Pred. No.:               5.56e-133     Length:        4527
Score:                   1438.00       Matches:       428
Percent Similarity:      44.54%        Conservative:  274
Best Local Similarity:   27.16%        Mismatches:    557
Query Match:             19.57%        Indels:        317
DB:                      13            Gaps:          46

US-10-020-095-4 (1-1428) x US-09-756-247-3 (1-4527)

Qy         20 LeuAlaValAlaProGly---------------ProArgPheLeuValThrAlaProGly 34
              ||||||::::::|||                  |||   :::|||||||||   |||
Db         28 TTGGCCCTATCACCAGCCATTGCAGAAGAACTTCCAAACTACCTGGTGACATTACCAGCC 87

Qy         35 IleIleArg---ProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHisCysPro 53
                 :::   |||            ::::::               |||:::      |||
Db         88 CGGCTAAATTTCCCC------------TCCGTTCAGAAGGTTTGTTTGGACCTGAGCCCT 135

Qy         54 SerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrValSerVal 73
                          ||||||      :::                :::   |||      :::
Db        136 GGGTACAGTGATGTTAAATTCACGGTTACTCTGGAGACCAAGGACAAGACCCAGAAGTTG 195

Qy         74 LeuGluAlaGluGlyValPheGluLysGlySerPheLysThrLeuThrLeuProSerLeu 93
              ||||||      |||:::   ::::::                         |||
Db        196 CTAGAATACTCTGGACTGAAGAAGAGGCACTTACATTGTATCTCCTTTCTTGTACCACCT 255

Qy         94 ProLeuAsnSerAlaAspGluIleTyrGluLeuArgValThrGlyArgThrGlnAspGlu 113
              |||            :::|||:::         :::||||||:::|||         :::
Db        256 CCTGCTGGTGGCACAGAAGAAGTGGCCACAATCCGGGTGTCGGGA---GTTGGAAATAAC 312

Qy        114 IleLeuPheSerAsnSerThrArgLeuSerPheGluThrLysArgIleSerValPheIle 133
              |||   |||               ::::::         :::   :::         |||:::
Db        313 ATCAGCTTTGAGGAGAAGAAAAAGGTTCTAATTCAGAGGCAGGGGAACGGCACCTTTGTA 372

Qy        134 GlnThrAspLysAlaLeuTyrLysProLysGlnGluValLysPheArgIleValThrLeu 153
              ||||||||||||      ||||||   |||   |||:::|||   ||||||||||||||:::
Db        373 CAGACTGACAAACCTCTCTACACCCCAGGGCAGCAAGTGTATTTCCGCATTGTCACCATG 432

Qy        154 PheSerAspPheLysProTyrLysThrSerLeuAsnIleLeu---IleLysAspProLys 172
                 |||:::|||   |||            :::::::::      ::::::||||||
Db        433 GATAGCAACTTCGTTCCAGTGAATGACAAGTACTCCATGGTGGAACTACAGGATCCAAAT 492

Qy        173 SerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAspLeuGlyValIleSerLysThr 192
              ||||||   |||   |||||||||                  :::   |||::::::   :::
Db        493 AGCAACAGGATTGCACAGTGGCTGGAAGTGGTACCTGAGCAAGGCATTGTAGACCTGTCC 552

Qy        193 PheGlnLeuSerSerHisProIleLeuGlyAspTrpSerIleGlnValGlnValAsnAsp 212
              |||||||||:::         :::||||||   ::::::         |||   |||   :::
Db        553 TTCCAACTGGCACCAGAGGCAATGCTGGGCACCTACACT------GTGGCAGTGGCTGAG 606

Qy        213 GlnThrTyrTyrGlnSerPheGlnValSerGluTyrValLeuProLysPheGluValThr 232
                       :::   :::|||   |||   |||||||||||||||||||||:::|||
Db        607 GGCAAGACCTTTGGTACTTTCAGTGTGGAGGAATATGTGCTGCCGAAGTTTAAGGTGGAA 666

Qy        233 LeuGlnThrProLeuTyrCysSerMetAsnSerLysHisLeuAsnGlyThrIleThrAla 252
              :::         |||         |||         :::               |||
Db        667 GTGGTGGAACCCAAGGAGTTATCAACGGTGCAGGAATCTTTCTTAGTAAAAATTTGTTGT 726

Qy        253 LysTyrThrTyrGlyLysProValLysGlyAspValThrLeuThrPheLeuProLeuSer 272
              :::||||||||||||||||||:::   |||   |||   ::::::
Db        727 AGGTACACCTATGGAAAGCCCATGCTAGGGGCAGTGCAGGTATCTGTGTGT--------- 777

Qy        273 PheTrpGlyLysLysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPhe 292
                       :::|||   |||                                  :::
Db        778 ---------CAGAAGGCAAAT---------------------------------ACTTAC 795

Qy        293 SerPheAsnAspGluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyr 312
                 :::   :::   |||   :::   :::   |||               ||||||
Db        796 TGGTATCGAGAGGTGGAACGGGAACAGCTTCCTGACAAATGCAGGAACCTCTCTGGACAG 855

Qy        313 LeuAsp------------LeuSerSerPro------------------------------ 318
                 |||            |||:::|||
Db        856 ACTGACAAAACAGGATGTTTCTCAGCACCTGTGGACATGGCCACCTTTGACCTCATTGGA 915

Qy        319 ------------GlyProValGluIleLeuThrThrValThrGluSerValThrGlyIle 334
                          :::   |||:::   ||||||   |||      ||||||:::
Db        916 TATGCGTACAGCCATCAAATCAATATTGTGGCTACTGTTGTGGAGGAAGGGACAGGTGTG 975

Qy        335 SerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePhe 354
                    |||   :::   |||::::::                     :::   |||
Db        976 GAGGCCAATGCCACTCAGAATATCTACATTTCTCCACAAATGGGATCAATGACCTTTGAA 1035

Qy        355 AspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLysValThrArg 374
              |||   :::               |||:::      |||:::         ::::::|||
```

```
Db   1036 GACACCAGCAATTTTTACCATCCAAATTTCCCCTTCAGTGGGAAGATAAGAGTTAGGGGC 1095

Qy    375 AlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle--------- 391
Db   1096 CATGATGACTCCTTCCTC---------------AAGAACCATCTAGTGTTTCTGGTGATT 1140

Qy    392 ------------------------------ThrValThrGlnArgAsn------------ 397
Db   1141 TATGGCACAAATGGAACCTTCAACCAGACCCTGGTTACTGATAACAATGGCCTAGCTCCC 1200

Qy    398 ---------TyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAla 414
Db   1201 TTTACCTTGGAGACATCCGGTTGGAATGGGACAGAC------------------------ 1236

Qy    415 ValGlnLysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGlu--------- 431
Db   1237 ---------------------GTTTCTCTGGAGGGAAAGTTTCAAATGGAAGACTTAGTA 1275

Qy    432 ---------------PheProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyr 446
Db   1276 TATAATCCGGAACAAGTGCCACGTTACTACCAAAATGCCTACCTGCACCTGCGACCCTTC 1335

Qy    447 PheLeuGlySerLysSerSerMetAlaValHisSerLeuPheLysSerProSerLysThr 466
Db   1336 TACAGCACAACCCGCAGCTTCCTTGGCATCCACCGGCTA--------------------- 1374

Qy    467 TyrIleGlnLeuLysThrArgAspGluAsnIleLysValGlySerProPheGluLeuVal 486
Db   1375 ---------------------AACGGCCCCTTGAAATGTGGCCAGCCCCAGGAAGTGCTG 1413

Qy    487 Val---------------------------SerGlyAsnLysArgLeuLysGluLeuSer 497
Db   1414 GTGGATTATTACATCGACCCGGCCGATGCAAGCCCTGACCAAGAGATC---AGCTTCTCC 1470

Qy    498 TyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGlnAsn------------ 513
Db   1471 TACTATTTAATAGGGAAAGGAAGTTTGGTGATGGAGGGGCAGAAACACCTGAACTCTAAG 1530

Qy    514 ------------------SerThrMetPheSerLeuThrProGluAsnSerTrpThrPro 527
Db   1531 AAGAAAGGACTGAAAGCCTCCTTCTCTCTCACTGACCTTCACTTCGAGACTGGCCCCT 1590

Qy    528 LysAlaCysValIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspValLeu 547
Db   1591 GATCCTTCCCTGGTGATCTATGCCATTTTTCCCAGTGGAGGTGTTGTAGCTGACAAAATT 1650

Qy    548 LysIleProValGlnLeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLys 567
Db   1651 CAGTTCTCAGTCGGGATGTGCTTTGACAATCAGGTTTCCCTTGGCTTCTCCCCCTCCCAG 1710

Qy    568 AlaGluProSerGluLysValSerLeuArgIleSerValThrGlnProAspSerIleVal 587
Db   1711 CAGCTTCCAGGAGCAGAAGTGGAGCTGCAGCTGCAGGCAGCT---CCCGGATCCCTGTGT 1767

Qy    588 GlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMet 607
Db   1768 GCGCTCCGGGCGGTGGATGAGAGTGTCTTACTGCTTAGGCCA------------------ 1809

Qy    608 GluAsnValValHisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPhe--- 626
Db   1810 ---------------GACAGAGAGCTGAGCAACCGCTCTGTCTAT---GGGATGTTTCCA 1851

Qy    626 ------------------------------------------------------------ 626
Db   1852 TTCTGGTATGGTCACTACCCCTATCAAGTGGCTGAGTATGATCAGTGTCCAGTGTCTGGC 1911

Qy    626 ------------------------------------------------------------ 626
Db   1912 CCATGGGACTTTCCTCAGCCCCTCATTGACCCAATGCCCCAAGGGCATTCGAGCCAGCGT 1971

Qy    627 --------------------------------MetAsnSerPheAlaValPheGlnGlu 635
Db   1972 TCCATTATCTGGAGGCCCTCGTTCTCTGAAGGCACGGACCTTTTCAGCTTTTTCCGGGAC 2031

Qy    636 CysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyr 655
Db   2032 GTGGGCCTGAAAATACTGTCCAATGCCAAAATCAAGAAGCCA---GTAGATTGCAGTCAC 2088

Qy    656 AspAsnAlaGluTyrAla---------------------------GluArgPheMetGlu 666
Db   2089 AGATCTCCAGAATACAGCACTGCTATGGGTGGCGGTGGTCATCCAGAGGCTTTTGAGTCA 2148

Qy    667 GluAsnGluGlyHisIleValAspIleHisAspPheSerLeuGlySerSerProHisVal 686
Db   2149 TCAACTCCTTTACATCAAGCAGAG------------------------GATTCTCAGGTC 2184

Qy    687 ArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyr 706
Db   2185 CGCCAGTACTTCCCAGAGACCTGGCTCTGGGATCTGTTTCCTATTGGTAACTCGGGGAAG 2244

Qy    707 GlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPheVal 726
Db   2245 GAGGCGGTCCACGTCACAGTTCCTGACGCCATCACCGAGTGGAAGGCGATGAGTTTCTGC 2304

Qy    727 IleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGln 746
Db   2305 ACTTCCCAGTCAAGAGGCTTCGGGCTTTCACCCACT---GTTGGACTAACTGCTTTCAAG 2361

Qy    747 ProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeu 766
Db   2362 CCGTTCTTTGTTGACCTGACTCTCCCTTACTCAGTAGTCCGTGGGGAATCCTTTCGTCTT 2421

Qy    767 GluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIleGluLys 786
Db   2422 ACTGCCACCATCTTCAATTACCTAAAGGATTGCATCAGGGTTCAGACTGACCTGGCTAAA 2481

Qy    787 SerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGln 806
Db   2482 TCGCATGAGTACCAGCTAGAATCATGGGCAGATTCT------------------CAGACC 2523

Qy    807 ThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgProThrHis 826
Db   2524 TCCAGTTGTCTCTGTGCTGATGACGCAAAAACCCACCACTGGAACATCACAGCTGTCAAA 2583

Qy    827 LeuGlyGluIleProIleThrValThrAlaLeu--------------------------- 837
Db   2584 TTGGGTCACATTAACTTTACTATTAGTACAAAGATTCTGGACAGCAATGAACCATGTGGG 2643

Qy    838 ---------------------SerProThrAlaSerAspAlaValThrGlnMetIleLeu 850
Db   2644 GGCCAGAAGGGGTTTGTTCCCCAAAAGGGCCGAAGTGACACGCTCATCAAGCCAGTTCTC 2703

Qy    851 ValLysAlaGluGly------IleGluLysSerTyrSerGlnSerIleLeuLeuAspLeu 868
Db   2704 GTCAAACCTGAGGGAGTCCTGGTGGAGAAGACACACAGC---------------TCATTG 2748

Qy    869 ThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsnThr 888
Db   2749 CTGTGCCCAAAAGGAAAGGTGGCATCTGAATCTGTCTCCCTGGAGCTCCCAGTGGACATT 2808

Qy    889 ValThrGlySerGluArgValGlnIleThrAlaIleGly--------------------- 901
Db   2809 GTTCCTGACTCGACCAAGGCTTATGTTACGGTTCTGGGTAAGCAGTTAGAGATTCTTGAC 2868

Qy    902 ----------------------------------------AspValLeuGlyProSer 907
Db   2869 TCAGAAAGGAAAAGGCGAATGGAGGCTGCAAAGGTGTGGAGAGACATTATGGGCACAGCC 2928

Qy    908 IleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIle 927
Db   2929 CTGCAGAACCTGGATGGTCTGGTGCAGATGCCCAGTGGCTGTGGCGAGCAGAACATGGTC 2988

Qy    928 AsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAsp 947
Db   2989 TTGTTTGCTCCCATCATCTATGTCTTGCAGTACCTGGAGAAGGCAGGGCTGCTGACGGAG 3048
```

```
Qy        948 AsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyr 967
                 :::::   :::|||:::   |||:::      |||||||||:::||||||:::|||
Db       3049 GAGATCAGGTCTCGGGCAGTGGGTTTCCTGGAAATAGGGTACCAGAAGGAGCTGATGTAC 3108

Qy        968 GlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrp 987
              :::      :::||||||:::||||||||||||         |||   :::|||:::||||||
Db       3109 AAACACAGCAATGGCTCATACAGTGCCTTTGGGGAGCGAGATGGAAATGGAAACACATGG 3168

Qy        988 LeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGln 1007
              |||:::|||||||||      :::||||||   :::|||      :::|||   ||||||
Db       3169 CTGACAGCGTTTGTCACAAAATGCTTTGGCCAAGCTCAGAAATTCATCTTCATTGATCCC 3228

Qy       1008 AsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPhe 1027
                    :::                  |||:::   |||:::|||      :::|||   :::
Db       3229 AAGAACATCCAGGATGCTCTCAAGTGGATGGCAGGAAACCAGCTCCCCAGTGGCTGCTAT 3288

Qy       1028 TrpAspProGlyArgValIleHisSerGluLeuGlnGlyGlyAsnLysSerProValThr 1047
                 :::   |||   ::::::|||:::   ::::::||||||            |||:::
Db       3289 GCCAACGTGGGAAATCTCCTTCACACAGCTATGAAGGGTGGTGTTGATGATGAGGTCTCC 3348

Qy       1048 LeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAsp 1067
              ||||||||||||:::         :::||||||      |||         :::
Db       3349 TTGACTGCGTATGTCACAGCTGCATTGCTGGAGATGGGAAAGGATGTAGATGACCCAATG 3408

Qy       1068 ValGlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyr 1087
              |||   :::   :::      |||::::::      :::         ::::::   |||
Db       3409 GTGAGTCAGGGTCTACGGTGTCTCAAGAATTCG---GCCACCTCCACGACCAACCTCTAC 3465

Qy       1088 ThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGlu--- 1106
              |||   ||||||:::   |||      |||      |||         :::
Db       3466 ACACAGGCCCTGTTGGCTTACATTTTCTCCCTGGCTGGGGAAATGGACATCAGAAACATT 3525

Qy       1107 AlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSer 1126
                 |||   |||      :::|||         |||         :::|||
Db       3526 CTCCTTAAACAGTTAGATCAACAGGCTATCATCTCAGGAGAATCCATTTACTGGAGCCAG 3585

Qy       1127 SerGluSerLysLeuSerAspSer------Trp---GlnProArgSerLeuAspIleGlu 1143
                 :::   |||::::::         |||   :::|||   ::::::|||:::|||
Db       3586 AAACCTACTCCATCATCGAACGCCAGCCCTTGGTCTGAGCCTGCGGCTGTAGATGTGGAA 3645

Qy       1144 ValAlaAlaTyrAlaLeuLeuSerHis---------------PheLeuGlnPheGlnThr 1158
              :::   |||||||||||||||:::                           |||   :::
Db       3646 CTCACAGCATATGCATTGTTGGCCCAGCTTACCAAGCCCAGCCTGACTCAAAAGGAGATA 3705

Qy       1159 SerGluGlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPhe 1178
              ::::::         |||:::   ||||||::::::|||   |||:::   ||||||||||
Db       3706 GCGAAGGCCACTAGCATAGTGGCTTGGTTGGCCAAGCAACACAATGCATATGGGGGCTTC 3765

Qy       1179 AlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla------Ala 1196
              :::||||||||||||||   ||||||||:::||||||:::::::::|||      |||
Db       3766 TCTTCTACTCAGGATACTGTAGTTGCTCTCCAAGCTCTTGCCAAATATGCCACTACCGCC 3825

Qy       1197 LeuMetAsnThrGluArgThrAsnIleGlnValThrValThr------------------ 1210
                 |||   :::|||      |||:::   |||   |||
Db       3826 TACATGCCATCTGAGGAGATCAACCTGGTTGTAAAATCCACTGAGAATTTCCAGCGCACA 3885

Qy       1211 ---GlyProSerSerProSerProLeuAlaValValGln------------------Pro 1223
                          |||   :::   |||      |||                  |||
Db       3886 TTCAACATACAGTCAGTTAACAGATTGGTATTTCAGCAGGATACCCTGCCCAATGTCCCT 3945

Qy       1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
                          :::   |||:::|||   |||         |||      :::
Db       3946 GGAATGTACACGTTGGAGGCCTCAGGCCAGGGCTGTGTCTATGTGCAGACGGTGTTGAGA 4005

Qy       1244 TyrAsnValLysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAla 1263
              ||||||:::                                      |||   :::
Db       4006 TACAATATTCTCCCTCCC--------------------------ACAAATATGAAGACC 4038

Qy       1264 PheAspLeuAspValAlaValLys------------GluAsnLysAspAspLeuAsnHis 1279
              |||   |||   |||   :::                |||
Db       4039 TTTAGTCTTAGTGTGGAAATAGGAAAAGCTAGATGTGAGCAACCGACTTCACCTCGATCC 4098

Qy       1280 ValAspLeuAsnValCysThrSerPheSerGlyPro---GlyArgSerGlyMetAlaLeu 1298
                 :::   |||   :::   ||||||:::   |||         |||   ||||||:::
Db       4099 TTGACTCTCACTATTCACACCAGTTATGTGGGGAGCCGTAGCTCTTCCAATATGGCTATT 4158

Qy       1299 MetGluValAsnLeuLeuSerGlyPhe------MetValProSerGluAlaIleSerLeu 1316
              :::||||||   :::||||||||||||         |||      :::   :::   |||
Db       4159 GTGGAAGTGAAGATGCTATCTGGGTTCAGTCCCATGGAGGGCACCAATCAGTTACTTCTC 4218

Qy       1317 SerGlu---ThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAsp 1335
                 :::         |||||||||||||||:::            ||||||:::||||||||
Db       4219 CAGCAACCCCTGGTGAAGAAGGTTGAATTTGGAACTGACACACTTAACATTTACTTGGAT 4278

Qy       1336 SerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSer 1355
                 :::   :::                              ::::::      |||:::
Db       4279 GAGCTCATTAAGAACACTCAGACTTACACCTTCACCATCAGCCAAAGTGTGCTGGTCACC 4338

Qy       1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaVal 1375
                 |||   :::   |||::::::   :::   |||||||||      |||
Db       4339 AACTTGAAACCAGCAACCATCAAGGTCTATGACTACTACCTACCA--------------- 4383

Qy       1376 ArgSerTyrAsnSerGluValLysLeuSer-SerCysAspLeuCys 1390
                                  |||||||||   |||      ||||||
Db       4384 ------------GGTTCTTTTAAATTATCTCAGTACACAATTGTGT 4417
```

Search completed: January 17, 2004, 08:49:35
Job time : 1238 secs

```
OM protein - protein search, using sw model

Run on:         January 15, 2004, 18:10:30 ; Search time 28 Seconds
                                             (without alignments)
                                             4904.603 Million cell updates/sec

Title:          US-10-020-095-4
Perfect score:  7348
Sequence:       1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters:      283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_76:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1854.5	25.2	1508	2	T27828	hypothetical prote
2	1850	25.2	1519	2	T27829	hypothetical prote
3	1638	22.3	1507	2	T18544	alpha-2-macroglobu
4	1491	20.3	1476	2	JC5143	alpha-macroglobuli
5	1455	19.8	1477	2	A29952	alpha-1 proteinase
6	1440	19.6	1474	1	MAHU	alpha-2-macroglobu
7	1424	19.4	1503	2	T43166	alpha-2-macroglobu
8	1417	19.3	1487	2	S15904	alpha-1 proteinase
9	1412.5	19.2	1472	2	A26122	alpha-2-macroglobu
10	1402	19.1	1451	2	B41185	alpha-2 macroglobu
11	1396.5	19.0	1482	2	S13495	pregnancy zone pro
12	1395.5	19.0	1476	2	A41185	alpha-2 macroglobu
13	1395	19.0	1473	2	A20872	ovostatin precurso
14	1384	18.8	1500	2	A42210	alpha-1-macroglobu
15	1374.5	18.7	1464	2	JC5144	murinoglobulin pre
16	1373.5	18.7	1495	2	S27001	alpha-2-macroglobu
17	1105.5	15.0	1699	2	T14074	complement compone
18	1079.5	14.7	1683	2	T30885	complement compone
19	1026.5	14.0	1673	2	I50806	complement compone
20	990	13.5	1652	2	I50711	complement C3 prec
21	987	13.4	1738	1	A24558	complement C4 prec
22	986.5	13.4	1736	2	A29176	sex-limited protei
23	982.5	13.4	1735	2	S54784	sex-limited protei
24	982.5	13.4	1744	1	C4HU	complement C4A pre
25	973.5	13.2	1651	1	C3NJ	complement C3 prec
26	966	13.1	1663	1	C3RT	complement C3 prec
27	963	13.1	785	2	A29953	alpha-1 proteinase
28	960	13.1	1620	2	I51339	complement compone
29	956	13.0	1663	1	C3MS	complement C3 prec
30	934	12.7	1663	1	C3HU	complement C3 prec
31	921	12.5	1676	1	C5HU	complement C5 prec
32	905	12.3	1680	1	C5MS	complement C5 prec
33	904	12.3	1666	1	C3GP	complement C3 prec
34	894	12.2	1642	2	I51018	cobra venom factor
35	857.5	11.7	1620	2	S21045	complement protein
36	736.5	10.0	609	2	A41081	alpha-1-inhibitor
37	694	9.4	1230	2	T30517	complement C3-Q2 -
38	688.5	9.4	1617	2	T28153	complement C4 - ch
39	467.5	6.4	1536	2	E72310	hypothetical prote
40	432.5	5.9	726	2	A27602	complement C3 - ra
41	290.5.	4.0	1892	2	C97804	hypothetical prote
42	287	3.9	1653	2	B91052	hypothetical prote
43	287	3.9	1653	2	F85896	hypothetical prote
44	283	3.9	1653	2	G65028	hypothetical prote
45	269.5	3.7	1644	2	AC0823	probable lipoprote

ALIGNMENTS

```
RESULT 1
T27828
hypothetical protein ZK337.1a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C;Accession: T27828
R;White, S.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z20426
A;Accession: T27828
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1508 <WIL>
A;Cross-references: EMBL:Z82090; PIDN:CAB05006.1; GSPDB:GN00019; CESP:ZK337.1a
A;Experimental source: clone ZK337
C;Genetics:
A;Gene: CESP:ZK337.1a
A;Map position: 1
A;Introns: 36/2; 116/3; 201/1; 267/2; 338/3; 413/3; 527/3; 556/3; 691/3; 714/1; 760/2; 8
C;Superfamily: alpha-2-macroglobulin

  Query Match             25.2%;  Score 1854.5;  DB 2;  Length 1508;
  Best Local Similarity   31.8%;  Pred. No. 6.3e-97;
  Matches  476;  Conservative  292;  Mismatches  562;  Indels  169;  Gaps   44;

Qy         15 VCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVL 74
              | |   | |     ::: || ::||            :    :      :| |::
Db         28 VSTTAAPVKPA-TYMLVAPAVVRPDQPFSVCMNLLKQATDEDMIVRIEVRTERNETIAAR 86

Qy         75 EAEGVFEKGSFKTLTLPSLPLNSAD--EIYELRVTGRTQD-EILFSNSTRLSFETKRISV 131
                : : |  :|::|  :|  |     : |:| : | | : |::| |   | :: | :||
Db         87 VISNL-KPGIAQTVSLSEMPAQSLTPRQSYKLYIRGETLNAELIFENENELKYDQKALSV 145

Qy        132 FIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISK 191
              |||||:|:|:|  |::| : : || |||  :  | | || ||| |  :    ||| |
Db        146 FIQTDRAIYRPASLVRYRAIVVKSDLKPYVGNATIKIFDPSRNLISQTIGVTLDRGVYSG 205

Qy        192 TFQLSSHPILGDWSIQVQ----VNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLN 247
               ||:   :|||| |:|:    | |::   || |  |||||||| ::|  : ::|    |:
Db        206 ELQLAEETLLGDWFIEVETSNGVQDKS---SFTVDTYVLPKFEVNIKTSSFITIND-DLS 261

Qy        248 GTITAKYTYGKPVKGDVTLTF-LPLSFW-----------GKKKN----ITKTFKIN--GS 289
               : ||||||| |     ::  ||  |             | ||         : :| |:| |
Db        262 VFVDAKYTYGKGVAGKAKVSLELPWHRWHAMVPTIIDENGVKKEEELMVERTVKLNRQGE 321

Qy        290 ANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDY 349
              |    |:::|:|          :  |      | : |: :|||  :| | || :  :  :  :
Db        322 AAVVFSNDELKR--------HKLLHEWGGGSIRIVASVTEDITEIERNATHQISTFREEV 373

Qy        350 IIEFFDYTTVLKPSLNFTATVKVTRADGN--QLTLEERRNNVVITVTQRNYTEYWSGSNS 407
               ::         || |  :   | : : |    : || :|    |    |  ||     ::  |
Db        374 KLDVEKQGDTFKPGLTYNVVVALKQMDDTPVKATLPKR----VQVSTFYNYP--YNHDTS 427
```

```
Qy     408 GNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK------ 461
Db     428 SLQEEKETKIVEVDAHGTSVLTLQPPINCTSARIEAH-YDIGGKDNFTATPIYSSLYVEA 486

Qy     462 --SPSKTYIQLKTRDEN-IKVGSPFELVVSGNKRLKELSYMVVSRGQLV-----AVGKQN 513
Db     487 AVSPTKSFLQLLADNEGAVDVGKSLSFSLKATQPLSTITYQVMSRSNIVVSQQMTVNSEH 546

Qy     514 STMFSLTPENSWTPKACVIVY-YIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSE 572
Db     547 ATI-SFPATANMAPKSRLIVYAIIESSQEVLVDALDFKVEGIFQNQVALSIDKQAVEPGQ 605

Qy     573 KVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMF------ 626
Db     606 NVKFKVT-SDKNSFVGLLVVDQSVLLLKTGNDITREKVEQDLENYDSNNVGGGFGGPRPW 664

Qy     627 -------------------MNSFAVFQECGLWVLTDANLTKD---------YIDGVYDNA 658
Db     665 EAIDRKKRSIWRPWWGIGGSDAQSIFSNAGLVVLTDALLYREPQREFMSVMMMDGAPGMA 724

Qy     659 EYAERFMEENEGHIVDIHDFSLGSS---PHVRKHFPETWIWLDTNMGYRIYQEFEVTVPD 715
Db     725 EAA--FAAPPMG----------GSSPPPPTVRKFFPHTWIWSDLN-STSGEVEMEIEAPD 771

Qy     716 SITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLK 775
Db     772 TITSWVASTFAINEENGLGVAPTTSKLRVFRPFFIQLNLPYAVRRGEKFALLVLVFNYME 831

Qy     776 DATEVKVII--EKSDKFDILMTSSEI---NATGHQQTLLVPSEDGAT---VLFPIRPTHL 827
Db     832 KEQDVTVTLKYDKDSGYDLLKKDGTVVRRDEVGQQNVRIVSVAGGGTSKAVYFPIVPSSI 891

Qy     828 GEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPN 887
Db     892 GEIPVHISAIASQGGDAVEMNLRVDPQGYKVDRNIPFVIDLNNNSSDFS-KNLELIWPND 950

Qy     888 TVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTD 947
Db     951 VVDGSQKARLDVIGDMMGPVLNNAHKLVQMPYGCGEQNMLNLVPNILVVKYLRATNRNES 1010

Qy     948 NLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQ 1007
Db    1011 QLETKAIKFIEQGIQRELTYKRADNSFSAFGDSDKAGSTWLTAFVVRSFHHAKQYAFVDP 1070

Qy    1008 NVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG-NKSPVTLTAYIVTSLLGYRKYQPNI 1066
Db    1071 NVISRAVAFLNSQQMESGAFAERGEVHHKDMQGGAQDGGVALTAFVLISIL-----ENGM 1125

Qy    1067 DVQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWR--AEQEGGMQFW 1124
Db    1126 ENGKAVTYLEKHLDEVSGNAYTMAVVAYALQLAKSKQAGKAFENLKKHKIVEKSGDVKFA 1185

Qy    1125 VSSE--SKLSDS----WQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGF 1178
Db    1186 SAQKKVEKLKESRAYMFQARPVDIETTSYAVLSYLAQNQTSESLSIIRWLVSQRNELGGF 1245

Qy    1179 ASTQDTTVALKALSEFAALMNTERTNIQVTV-TGPSSPS------PLAVVQPM------- 1224
Db    1246 TSTQDTVMALQALSSYAAVTYSDKHTSQVTILNGKHTHSFDINIRNAIVLQSYQLSSLND 1305

Qy    1225 AVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNV 1284
Db    1306 AVSINANGTGVVFAQLSYSY---------YRDSLNDDAPFFCSQEIKEIRAG-NRLQLDL 1355

Qy    1285 CTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVK-----KVEYDHGKLNLYLDSVNE 1339
Db    1356 CCNYTRPGKSNMALAEIDALSGYRFDAEQVHTLTSIEDLQRVEMEKDDTKMNVYFNPLGG 1415

Qy    1340 TQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRP 1398
Db    1416 RPVCLSLYSDVTYQVADQKPANFRLVDYYDPEEQLKMTYAAKQTRSLQEKCG--EDCWP 1472
```

RESULT 2
T27829
hypothetical protein ZK337.1b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C;Accession: T27829
R;White, S.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z20426
A;Accession: T27829
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1519 <WIL>
A;Cross-references: EMBL:Z82090; PIDN:CAB05007.1; GSPDB:GN00019; CESP:ZK337.1b
A;Experimental source: clone ZK337
C;Genetics:
A;Gene: CESP:ZK337.1b
A;Map position: 1
A;Introns: 36/2; 116/3; 201/1; 267/2; 338/3; 413/3; 527/3; 556/3; 691/3; 714/1; 771/2; 8
C;Superfamily: alpha-2-macroglobulin

```
  Query Match             25.2%;  Score 1850;  DB 2;  Length 1519;
  Best Local Similarity   31.6%;  Pred. No. 1.1e-96;
  Matches  477;  Conservative  290;  Mismatches  563;  Indels  180;  Gaps   44;

Qy      15 VCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVL 74
Db      28 VSTTAAPVKPA-TYMLVAPAVVRPDQPFSVCMNLLKQATDEDMIVRIEVRTERNETIAAR 86

Qy      75 EAEGVFEKGSFKTLTLPSLPLNSAD--EIYELRVTGRTQD-EILFSNSTRLSFETKRISV 131
Db      87 VISNL-KPGIAQTVSLSEMPAQSLTPRQSYKLYIRGETLNAELIFENENELKYDQKALSV 145

Qy     132 FIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISK 191
Db     146 FIQTDRAIYRPASLVRYRAIVVKSDLKPYVGNATIKIFDPSRNLISQTIGVTLDRGVYSG 205

Qy     192 TFQLSSHPILGDWSIQVQ----VNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLN 247
Db     206 ELQLAEETLLGDWFIEVETSNGVQDKS---SFTVDTYVLPKFEVNIKTSSFITIND-DLS 261

Qy     248 GTITAKYTYGKPVKGDVTLTF-LPLSFW-----------GKKKN----ITKTFKIN--GS 289
Db     262 VFVDAKYTYGKGVAGKAKVSLELPWHRWHAMVPTIIDENGVKKEEELMVERTVKLNRQGE 321

Qy     290 ANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDY 349
Db     322 AAVVFSNDELKR--------HKLLHEWGGGSIRIVASVTEDITEIERNATHQISTFREEV 373

Qy     350 IIEFFDYTTVLKPSLNFTATVKVTRADGN--QLTLEERRNNVVITVTQRNYTEYWSGSNS 407
Db     374 KLDVEKQGDTFKPGLTYNVVVALKQMDDTPVKATLPKR----VQVSTFYNYP--YNHDTS 427

Qy     408 GNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK------ 461
Db     428 SLQEEKETKIVEVDAHGTSVLTLQPPINCTSARIEAH-YDIGGKDNFTATPIYSSLYVEA 486

Qy     462 --SPSKTYIQLKTRDEN-IKVGSPFELVVSGNKRLKELSYMVVSRGQLV-----AVGKQN 513
Db     487 AVSPTKSFLQLLADNEGAVDVGKSLSFSLKATQPLSTITYQVMSRSNIVVSQQMTVNSEH 546

Qy     514 STMFSLTPENSWTPKACVIVY-YIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSE 572
Db     547 ATI-SFPATANMAPKSRLIVYAIIESSQEVLVDALDFKVEGIFQNQVALSIDKQAVEPGQ 605

Qy     573 KVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMF------ 626
Db     606 NVKFKVT-SDKNSFVGLLVVDQSVLLLKTGNDITREKVEQDLENYDSNNVGGGFGGPRPW 664

Qy     627 -------------------MNSFAVFQECGLWVLTDA--------------------NLT 647
Db     665 EAIDRKKRSIWRPWWGIGGSDAQSIFSNAGLVVLTDALLYREPQREFMSERRLNTPGGLT 724
```

```
Qy        648 KDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGSS---PHVRKHFPETWIWLDTNMGYR 704
Db        725 VMMMDGAPGMAEAA--FAAPPMG----------GSSPPPPTVRKFFPHTWIWSDLN-STS 771
Qy        705 IYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEF 764
Db        772 GEVEMEIEAPDTITSWVASTFAINEENGLGVAPTTSKLRVFRPFFIQLNLPYAVRRGEKF 831
Qy        765 ALEITIFNYLKDATEVKVII--EKSDKFDILMTSSEI---NATGHQQTLLVPSEDGAT-- 817
Db        832 ALLVLVFNYMEKEQDVTVTLKYDKDSGYDLLKKDGTVVRRDEVGQQNVRIVSVAGGGTSK 891
Qy        818 -VLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQST 876
Db        892 AVYFPIVPSSIGEIPVHISAIASQGGDAVEMNLRVDPQGYKVDRNIPFVIDLNNNSSDFS 951
Qy        877 LKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYIL 936
Db        952 -KNLELIWPNDVVDGSQKARLDVIGDMMGPVLNNAHKLVQMPYGCGEQNMLNLVPNILVV 1010
Qy        937 DYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF 996
Db       1011 KYLRATNRNESQLETKAIKFIEQGIQRELTYKRADNSFSAFGDSDKAGSTWLTAFVVRSF 1070
Qy        997 LEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG-NKSPVTLTAYIVTS 1055
Db       1071 HHAKQYAFVDPNVISRAVAFLNSQQMESGAFAERGEVHHKDMQGGAQDGGVALTAFVLIS 1130
Qy       1056 LLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWR- 1114
Db       1131 IL-----ENGMENGKAVTYLEKHLDEVSGNAYTMAVVAYALQLAKSKQAGKAFENLKKHK 1185
Qy       1115 -AEQEGGMQFWVSSE--SKLSDS----WQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRW 1167
Db       1186 IVEKSGDVKFASAQKKVEKLKESRAYMFQARPVDIETTSYAVLSYLAQNQTSESLSIIRW 1245
Qy       1168 LSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTV-TGPSSPS------PLAV 1220
Db       1246 LVSQRNELGGFTSTQDTVMALQALSSYAAVTYSDKHTSQVTILNGKHTHSFDINIRNAIV 1305
Qy       1221 VQPM-------AVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKEN 1273
Db       1306 LQSYQLSSLNDAVSINANGTGVVFAQLSYSY---------YRDSLNDDAPFFCSQEIKEI 1356
Qy       1274 KDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVK-----KVEYDHG 1328
Db       1357 RAG-NRLQLDLCCNYTRPGKSNMALAEIDALSGYRFDAEQVHTLTSIEDLQRVEMEKDDT 1415
Qy       1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCD 1388
Db       1416 KMNVYFNPLGGRPVCLSLYSDVTYQVADQKPANFRLVDYYDPEEQLKMTYAAKQTRSLQE 1475
Qy       1389 LCSDVQGCRP 1398
Db       1476 KCG--EDCWP 1483
```

RESULT 3
T18544
alpha-2-macroglobulin - Limulus sp.
C;Species: Limulus sp.
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000
C;Accession: T18544
R;Iwaki, D.; Kawabata, S.; Miura, Y.; Kato, A.; Armstrong, P.B.; Quigley, J.P.; Nielsen,
Eur. J. Biochem. 242, 822-831, 1996
A;Title: Molecular cloning of Limulus alpha-2-macroglobulin.
A;Reference number: Z18961; MUID:97175055; PMID:9022715
A;Accession: T18544
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1507 <IWA>
A;Cross-references: EMBL:D83196; NID:d1096438; PID:d1020631; PIDN:BAA19844.1
A;Experimental source: adult
C;Superfamily: alpha-2-macroglobulin

```
  Query Match             22.3%;  Score 1638;  DB 2;  Length 1507;
  Best Local Similarity   28.9%;  Pred. No. 1.3e-84;
  Matches  449;  Conservative  270;  Mismatches  559;  Indels  278;  Gaps   43;

Qy         28 FLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVS---------VLEAEG 78
Db         29 FILTAPKSLTPGKSNILNLHLFD-----------IKTNGFLRIGVKDQDDGNVVAETEV 76
Qy         79 VFEK---GSFKTLTLPS-----LPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRIS 130
Db         77 SFNKDNPSSSIQLTIPSGVEVKRPKLYANGSY----SSPSSNDFFFEKD--INMHKDKLI 130
Qy        131 VFIQTDKALYKPKQEVKFRIVTLFSDFK--PYKTSLNILIKDPKSNLIQQWLSQQSDLGV 188
Db        131 VFVQTDKPLYKPGQTVKVRILPTTPDLKLVPKETIGSFQIENPDGIVLGYWPMLSFAEGI 190
Qy        189 ISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLNG 248
Db        191 AQFELALPDEPTYGMWRIKGNIEDTEIYENFEVKEYVLPKFEVKITPPSYLLTNADSITW 250
Qy        249 TITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFK-----------INGSANFSFNDE 297
Db        251 KICAQYTYGQPVEGTFVAETNVVKYNWEKEGVPVIHKEGLIDGCLDVTVNSSA-LGFNEQ 309
Qy        298 EMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYT 357
Db        310 RL-----------------SYRAVNMFAEVTEKGTGIKMNATDSIYRTSNPLNIMYLEPT 352
Qy        358 T---VLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEA 414
Db        353 SGKGYLKPGLPFYGKLKVEKPDGTPAPGEQ------IELCRFADRERWNRKRWLEEKIRA 406
Qy        415 VQKINYTVPQSGTFKIEF-PILEDSSELQLKAYFL-------GSKSSMAVHSL----FKS 462
Db        407 CKE--FTSDEAGIIKFTVPPQTPDITSFRFKAKALQYGKKDGDNKLNQPQHSFTVSSWYS 464
Qy        463 PSKTYIQLKTRDENIKVGSPFEL---VVSGNKRLKELSYMVVSRGQLVAVGK-------- 511
Db        465 PSGSHLQLEPITEEIECGKPLTVKFKYTTGEEKKQKFYYQIMARNFIVDTGSFEHEFLLS 524
Qy        512 -------------QNSTMFSLTPEN--SW----------------------TPKACVIV 533
Db        525 EDKSGLTDETYLPIDVTALSLNPPNEPEWENNVIVPPHIGETSLTLIPSFEMNPSAKILV 584
Qy        534 YYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVD 593
Db        585 FYVREDGETVADSTKITVKKCLRNKVGLKFGEEKVLPGASSTLQLTAS-PYSICGIGAVD 643
Qy        594 KSVNLMNASNDITMENVVHEL------------------------------ELYNTGYY 622
Db        644 KSVHILSSDNRITEEEVFNKLGGHDYYWPKQATSDYKYCEDYKFKQTEGEHEGSFSSGFT 703
Qy        623 LGMFMNSFAVFQECGLWVLTDANL------TKDYIDGVYDNAEYAERFMEENEGHIV--- 673
Db        704 STNYLDSITAFDEAGLVVISDMELETRPCKPSGFEDGGRPCPQYDVAFAAPQAANRIGGG 763
Qy        674 -DIHDFSLG-----SSP--HVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGF 725
Db        764 GEAGGFGGGIRKKTNKPVVEIRTYFPETWLWELQNIGATGELSLKRDIPHTITEWVGSAI 823
Qy        726 VISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIE 785
Db        824 CISEETGLGVSEA-ATVKGFQPFFVSFTLPYSVIRGEKVPIIVTVFNYLSECLPIKLSLE 882
Qy        786 KSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS------- 838
Db        883 QSDKF-------EMQNDTNSYTSCVCGGKSDTTRWMIKPRSLGQVNLTVYGASLPNEAIC 935
Qy        839 -------PTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTG 891
```

```
Db        936 GNQDYSTVTARDAATRQLLVEPEGFPKEDTWSTFACPKDQNGKFT-ATSDLLLPEDLVED 994

Qy        892 SERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKE 951

Db        995 SARGYVSITGDLMGPAIKNLDHLVRLPTGCGEQNMVKFVPNIFVLDYLTATGSITDSIKE 1054

Qy        952 KALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLH 1011

Db       1055 KALNNMRKGYARQQNYRHPDGSYSAFGNRDKQGNLFLTAFVYRSFAQAERFILINKNKLN 1114

Qy       1012 RTYTWLKGHQKSNGEFWDPGRVIHSELQGG----NKSPVTLTAYIVTSLL--GYRKYQPN 1065

Db       1115 ETENWILNRQRSNGCFRKIGKLFNSALKGGISSNDETPAPLTAYVLISLLEAGYKN---- 1170

Qy       1066 IDVQESIHFLESEFSRGIS--------DNYTLALITYALSSVGSPKAKEALNMLTWRAEQ 1117

Db       1171 ----------ETVIDQGISCLEALSNPSTYSLALFAYATSLAGHPSAKDYLAKLEERAIT 1220

Qy       1118 EGGMQFWVSSESKLSDSWQPRSLDIEVAAYA---LLSHFLQFQTSEGIPIMRWLSRQRNS 1174

Db       1221 EGGKTFWKSPSS--GRYYWGNSIGVEIAGYAVLTLLQHGGASNLAKVTPIIRWLAKQQNY 1278

Qy       1175 LGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222

Db       1279 RGGFYSTQDTVIALQAMSKFATIIYKDELDLEVGVESSGFEKKIMLTKDNSILMQTFRLQ 1338

Qy       1223 --PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKEN--KDDLN 1278

Db       1339 TVPSPVDFEATGSGCGLVQTSLRYNVNTP---------PPRKGFHLEVTVKRGLYRDCIN 1389

Qy       1279 HVDLNVCTSFSGPGR-SGMALMEVNLLSGFMVPSEAIS-----LSETVKKVEYDHGKLNL 1332

Db       1390 -AHIATCVKYDGKGGVSNMAVLEMKMVSGWIPDEESIKNIVDREELNLRRYEVDGNQLNL 1448

Qy       1333 YLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCD 1388

Db       1449 YFSELTDQNLCFNFWLEQDIEVQETKPATIRLYDYYELEQEVVTSYSID---ENCE 1501
```

RESULT 4
JC5143
alpha-macroglobulin precursor - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C;Accession: JC5143
R;Iwasaki, H.; Suzuki, Y.; Sinohara, H.
J. Biochem. 120, 1167-1175, 1996
A;Title: Cloning and sequencing of cDNAs encoding plasma alpha-macroglobulin and murinog
A;Reference number: JC5143; MUID:97164019; PMID:9010766
A;Contents: liver
A;Accession: JC5143
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1476 <IWA>
A;Cross-references: DDBJ:D84338; NID:g1805591; PIDN:BAA12316.1; PID:g1304084
C;Comment: This protein belongs to alpha 2-macroglobulin family.
C;Superfamily: alpha-2-macroglobulin
F;1-23/Domain: signal sequence #status predicted <SIG>

```
  Query Match             20.3%;  Score 1491;  DB 2;  Length 1476;
  Best Local Similarity   28.8%;  Pred. No. 2.8e-76;
  Matches  451;  Conservative  292;  Mismatches  524;  Indels  300;  Gaps   60;

Qy          4 PPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAEL- 62

Db         11 PVLL----LLILLAADASISGKPQYMVLVPSLLHSGTPEKICL-LLTQLNETVTVKASLD 65

Qy         63 -LKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTR 121

Db         66 TIRENGSLFMNMVAEKDLFQCVAF---TVPQSPYPEAVMFLTVEVEGPTHG---FRSRKT 119

Qy        122 LSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLN-----ILIKDPKSNLI 176

Db        120 VLVKSKDSLVFVQTDKPIYKPGQTVKARVVSLDENFRP----LNELFPLIFIQDPKGNRV 175

Qy        177 QQWLSQQSDLGVISKTFQLSSHPILGDWSIQV-QVNDQTYYQSFQVSEYVLPKFEVTLQT 235

Db        176 MQWQNLKLERGLTQLSFPLSSEPLLGSYSVVVHKESGGRMHHSFTVEEFVLPKFEVQVSM 235

Qy        236 PLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFL-----PLSFWGKKKNITKTF--KING 288

Db        236 PKKITILEQEFTVSVCGRYTYGKPVPGNITMSICRNYNNPSACLSEE---SRAFCKKYNQ 292

Qy        289 SAN--------FSFNDEEMK------------NVMDSSNGLSEYLDLSSPGPVEILTTVT 328

Db        293 QLNSQGCFIQQVKTNDFQLRRKEYEMRLRVEAKIREEGTG----VQLTGTGFSEITATIT 348

Qy        329 ESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNN 388

Db        349 K--------------------LSFVKVDSYVRPGVPFFGQVRL--VDGKNVPMPHKM-- 383

Qy        389 VVITVTQRNYTEYWSGSNSGNQKMEAVQ-KINYTVPQSGTFKIEFPILEDSS----ELQL 443

Db        384 ITITASEANY-----HSNATTDENGLVQFSINTTNMIGTSLNIQVKHKDSTNCYDYQWLL 438

Qy        444 KAYFLGSKSSMAVHSLFKSPSKTYIQLKT--------RDENIKVGSPFELVVSGNKRLKE 495

Db        439 EANEGASHTANAVFSL----SRSFVHLEPQLGKLPCHQTQTFKA----HYILKG-QELKE 489

Qy        496 L--SYMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDGEII 543

Db        490 LVFYYVIMAKGGIVQSGTYVLSVEQGNTKGHFSVSVPVESDLAPVARVLIYAILPSGEII 549

Qy        544 SDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNL----- 598

Db        550 ADSAKYNVENCLDNKVNLSFSEGQSLPASKTHLRVTAS-PQSLCALRAVDQSVLLRKPEA 608

Qy        599 -MNASN--------DIT---------MENVVHELELYNTGYYLGMF---------MNSFA 631

Db        609 VLSASSVYALLPVKDLTGFPGLLGQQEENDGECVSLYNT-YIDGILYSPEPNINEKDMYG 667

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGSSPH------ 685

Db        668 FLKDMGLKVFTNTKIQKPQLCAHVQKFEVPTMAYSYSESS-------SFRSGPRRVPAVG 720

Qy        686 ------------VRKHFPETWIW---LDTNMGYRIYQEFEVTVPDSITSWVATGFVISED 730

Db        721 IAATYSEPPKETVRTYSPETWIWDLKVTDSSG---VAEVEVTVPDTITEWKAGAFCLSND 777

Qy        731 LGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKF 790

Db        778 TGLGLSPT-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPDCIRISVHLEASPKF 836

Qy        791 DILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTA-------------- 836

Db        837 LAEPKAKE------QESYCVCGNERQTVSWVVTPKSLGNVNFTVSAEALESSELCGNEKT 890

Qy        837 LSPT--ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSER 894

Db        891 VVPTYGKKDTIIKPLLVEPEGIEKEETWTSLIRVSDTTVS---EKLHLELPSNVIQDSAR 947

Qy        895 VQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKAL 954

Db        948 ATVSILGDILGSAMQNIQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPDIKSKAI 1007

Qy        955 SFMRQGYQRELLYQREDGSFSAFG-NY-DPSGSTWLSAFVLRCFLEADPYIDIDQNVLHR 1012

Db       1008 SYLSTGYQRQLNYKHRDGSYSTFGENYRGGQGNTWLTAFVLKTFSQARKYIFIDEAHITQ 1067

Qy       1013 TYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESI 1072

Db       1068 ALSWLSQKQKDNGCFWSSGSLLNNAIKGGVEDEISLSAYITIALLEMSLPDTHPVVRNAL 1127

Qy       1073 HFLES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFWV 1125

Db       1128 FCLESAWKSAKEGTHG-SHVYTKALLAYAFALAGNQERKKEILKSLEDEGVKEDNSLHWA 1186
```

```
Qy       1126 SSES-KLSDSW----QPRSLDIEVAAYALLSHFLQFQTSEGIP----------IMRWLSR 1170
Db       1187 RPQKPKVSEGFLFKSQAPSAEVEMTSYVLLAYL----TARPAPTPEDLTSATDIVNWVTK 1242

Qy       1171 QRNSLGGFASTQDTTVALKALSEFAA--LMNTERTNIQVTVTGPSSPSP--------LAV 1220
Db       1243 QQNSHGGYSSTQDTVVALHALSKYAAATFTRTEKA-AQVTIKSSGTFSTNFEVNHNNRLL 1301

Qy       1221 VQPMAV-------NISANGFGFAICQLNVVYNV-----------------KASGSSRRRR 1256
Db       1302 LQQVSLPTVSDSYTITVTGEGNVYLQTSLKYNVPSEKGTFPFALEAETVPQACDGPKAHT 1361

Qy       1257 SIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGF--MVPS-EA 1313
Db       1362 SFQ--------------------ISLNVSYIGSRP-VSNMAIVDVKMVSGFIPLKPTVKN 1400

Qy       1314 ISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQ 1373
Db       1401 LEKSEHISRTEVSNNHVLIYLDKVSNQTLSLSFFVVQDIEVRDLKPAIIKVYDYYETNEF 1460

Qy       1374 AVRSYNS 1380
Db       1461 AIAEYHA 1467
```

RESULT 5
A29952
alpha-1 proteinase inhibitor III precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 16-Jul-1999
C;Accession: A29952; A27199; A34278
R;Braciak, T.A.; Northemann, W.; Hudson, G.O.; Shiels, B.R.; Gehring, M.R.; Fey, G.H.
J. Biol. Chem. 263, 3999-4012, 1988
A;Title: Sequence and acute phase regulation of rat alpha-1-inhibitor III messenger RNA.
A;Reference number: A92709; MUID:88153707; PMID:2831216
A;Accession: A29952
A;Molecule type: mRNA
A;Residues: 1-1477 <BRA>
A;Cross-references: GB:J03552; NID:g202571; PIDN:AAA40628.1; PID:g202572
R;Schweizer, M.; Takabayashi, K.; Geiger, T.; Laux, T.; Biermann, G.; Buhler, J.M.; Gaut
Eur. J. Biochem. 164, 375-381, 1987
A;Title: Identification and sequencing of cDNA clones for the rodent negative acute-phas
A;Reference number: A27199; MUID:87190405; PMID:2436907
A;Accession: A27199
A;Molecule type: mRNA
A;Residues: 1253-1477 <SCH>
A;Cross-references: GB:M28297; NID:g204930; PIDN:AAA63493.1; PID:g204931
A;Experimental source: liver, clone pI3cDNA6
R;Enghild, J.J.; Salvesen, G.; Thogersen, I.B.; Pizzo, S.V.
J. Biol. Chem. 264, 11428-11435, 1989
A;Title: Proteinase binding and inhibition by the monomeric alpha-macroglobulin rat alph
A;Reference number: A34278; MUID:89291900; PMID:2472396
A;Accession: A34278
A;Molecule type: protein
A;Residues: 673-676,'T',678-723 <ENG>
C;Superfamily: alpha-2-macroglobulin
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-1477/Product: alpha-1-inhibitor III #status predicted <MAT>

```
  Query Match             19.8%;  Score 1455;  DB 2;  Length 1477;
  Best Local Similarity   29.4%;  Pred. No. 3.1e-74;
  Matches  457;  Conservative  237;  Mismatches  578;  Indels  284;  Gaps   50;

Qy         13 LCVCTAALAVAPGPRFLVTAPGIIRPGGN----VTIGVELLEHCPSQVTVKAELLKTASN 68
Db          9 LCLFSALLAFLPFASLL---------NGNSKYMVLVPSQLYTETPEKICLHLYHLNETVT 59

Qy         69 LTVSVLEAEG--------VFEKGSFK--TLTLPSLPLNSADEIYELRVTGRTQDEILFSN 118
Db         60 VTASLISQRGTRKLFDELVVDKDLFHCVSFTIPRLPSSEEEESLDINIEGAKHK---FSE 116

Qy        119 STRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-----IKDPKS 173
Db        117 RRVVLVKNKESVVFVQTDKPMYKPGQSVKFRVVSMDKNLHP----LNELFPLAYIEDPKM 172

Qy        174 NLIQQWLSQQSDLGVISKTFQLSSHPILGDWSI----QVQVNDQTYYQSFQVSEYVLPKF 229
Db        173 NRIMQWQDVKTENGLKQLSFSLSAEPIQGPYKIVILKQSGVKEE---HSFTVMEFVLPRF 229

Qy        230 EVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGS 289
Db        230 GVDVKVPNAISVYDEIINVTACATYTYGKPVPGHVKISL-----------------CHGN 272

Qy        290 ANFSFNDEEMKNVMDS---SNGLS------------EYLDLSSPGPVEILTTVTESVTG 333
Db        273 PTFSSETKSGCKEEDSRLDNNGCSTQEVNITEFQLKENYLKMHQ--AFHVNATVTEEGTG 330

Qy        334 ISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITV 393
Db        331 SEFSGSGRIEVERTRNKFLFLKADSHFRHGIPFFVKVRLVDIKGDPIPNEQ----VLIKA 386

Qy        394 TQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSS 453
Db        387 RDAGYTNATTTDQHGLAK---------------FSIDTNGISDYS-LNIKVYH--KEES 427

Qy        454 MAVHS---------------LFKSPSKTYIQLKTR------DENIKVGSPFEL--VVSGN 490
Db        428 SCIHSSCTAERHAEAHHTAYAVYSLSKSYIYLDTEAGVLPCNQIHTVQAHFILKGQVLGV 487

Qy        491 KRLKELSYMVVSRGQLVAVGK-----------QNSTMFSLTPENSWTPKACVIVYYIED 538
Db        488 LQQIVFHYLVMAQGSILQTGNHTHQVEPGESQVQGNFALEIPVEFSMVPVAKMLIYTILP 547

Qy        539 DGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNL 598
Db        548 DGEVIADSVKFQVEKCLRNKVHLSFSPSQSLPASQTHMRVTAS-PQSLCGLRAVDQSVLL 606

Qy        599 -----------------MNASNDITMENVVHELELYNTGYY-LGMFMNSFAVFQECGLWV 640
Db        607 QKPEAELSPSLIYDLPGMQDSNFIASSNDPFEDEDYCLMYQPIAREKDVYRYVRETGLMA 666

Qy        641 LTDANL-TKDYIDGVYDNAEYA----------ERFMEENEGHIVDIHDFSLGSSP----- 684
Db        667 FTNLKIKLPTYCNTDYDMVPLAVPAVALDSSTDRGMYESL-PVVAVKS-PLPQEPPRKDP 724

Qy        685 --------HVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLT 736
Db        725 PPKDPVIETIRNYFPETWIWDLVTVNSSGVTELEMTVPDTITEWKAGALCLSNDTGLGLS 784

Qy        737 TTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTS 796
Db        785 SV-ASFQAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPTSLPMAVLLEASPDF------ 837

Qy        797 SEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----PTAS--------- 842
Db        838 TAVPVENNQDSYCLGANGRHTSSWLVTPKSLGNVNFSVSAEARQSPGPCGSEVATVPETG 897

Qy        843 --DAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
Db        898 RKDTVVKVLIVEPEGIKKEHTFSSLLCASDAELS---ETLSLLLPPTVVKDSARAHFSVM 954

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
Db        955 GDILSSAIKNTQNLIQMPYGCGEQNMVLFAPNIYVLKYLNETQQLTEKIKSKALGYLRAG 1014

Qy        961 YQRELLYQREDGSFSAFGNYD--PSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLK 1018
Db       1015 YQRELNYKHKDGSYSAFGDHNGQGQGNTWLTAFVLKSFAQARAFIFIDESHITDAFTWLS 1074

Qy       1019 GHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNID--VQESIHFLE 1076
Db       1075 KQQKDSGCFRSSGSLLNNAMKGGVDDEITLSAYITMALL--ESSLPDTDPVVSKALSCLE 1132

Qy       1077 SEFSR-----GISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW------ 1124
```

```
Db       1133 SSWENIEQGGNGSFVYTKALMAYAFALAGNQEKRNEILKSLDKEAIKEDNSIHWERPQKP 1192

Qy       1125 VSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIP---------IMRWLSRQRNS 1174
               ||  |      |  |  | ::|::|| :|:      |::   |           ::||::|:||
Db       1193 TKSEGYLYTP-QASSAEVEMSAYVVLARL----TAQPAPSPEDLALSMGTIKWLTKQQNS 1247

Qy       1175 LGGFASTQDTTVALKALSEF-AALMNTERTNIQVTVTGPSSPSPLAVVQ---------- 1222
               |||:||||| ||| |||:: ||   :   :     |||     | |      |
Db       1248 YGGFSSTQDTVVALDALSKYGAATFSKSQKTPSVTVQSSGSFSQKFQVDKSNRLLLQQVS 1307

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK----ENK 1274
                   |     :|  :|  |      |   : |||                :  |  || |  |          |
Db       1308 LPYIPGNYTVSVSGEGCVYAQTTLRYNVPLE---------KQQPAFALKVQTVPLTCNNP 1358

Qy       1275 DDLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGFMVPSEAISLSETVKKV---------E 1324
                 |    :::   |:  |      |  |  :  : :|||     ||  ||||:             |
Db       1359 KGQNSFQISLEISYMGSRPASNMVIADVKMLSGF------IPLKPTVKKLERLGHVSRTE 1412

Qy       1325 YDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
                   : ||||  |          :       ::   |  |  |  |  |  : ||||     |     |:|
Db       1413 VTTNNVLLYLDQVTNQTLSFSFIIQQDIPVKNLQPAIVKVYDYYETDEVAFAEYSS 1468
```

RESULT 6
MAHU
alpha-2-macroglobulin precursor [validated] - human
N;Alternate names: alpha-2M
C;Species: Homo sapiens (man)
C;Date: 05-Apr-1983 #sequence_revision 30-Jun-1987 #text_change 08-Dec-2000
C;Accession: A94033; I39375; S09107; JN0262; A92486; S66634; A01256
R;Kan, C.C.; Solomon, E.; Belt, K.T.; Chain, A.C.; Hiorns, L.R.; Fey, G.
Proc. Natl. Acad. Sci. U.S.A. 82, 2282-2286, 1985
A;Title: Nucleotide sequence of cDNA encoding human alpha-2-macroglobulin and assignment
A;Reference number: A94033; MUID:85190481; PMID:2581245
A;Accession: A94033
A;Molecule type: mRNA
A;Residues: 1-1474 <KAN>
A;Cross-references: GB:M11313; NID:g177869; PIDN:AAA51551.1; PID:g177870
A;Note: hydrolysis of the thiolester bond during amino acid sequencing of the mature pro
lymorphism
R;Bell, G.I.; Rall, L.B.; Sanchez-Pescador, R.; Merryweather, J.P.; Scott, J.; Eddy, R.L
Somat. Cell Mol. Genet. 11, 285-289, 1985
A;Title: Human alpha 2-macroglobulin gene is located on chromosome 12.
A;Reference number: I39375; MUID:85219061; PMID:2408344
A;Accession: I39375
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 832-999,'I',1001-1147,'D',1149-1194,'D',1196-1474 <BEL>
A;Cross-references: GB:M36501; NID:g177871; PIDN:AAA51552.1; PID:g177872
R;Marynen, P.; Devriendt, K.; van den Berghe, H.; Cassiman, J.J.
FEBS Lett. 262, 349-352, 1990
A;Title: A genetic polymorphism in a functional domain of human pregnancy zone protein:
n.
A;Reference number: S09106; MUID:90242963; PMID:1692292
A;Accession: S09107
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 672-747 <MAR>
R;Matthijs, G.; Devriendt, K.; Cassiman, J.J.; Van Den Berghe, H.; Marynen, P.
Biochem. Biophys. Res. Commun. 184, 596-603, 1992
A;Title: Structure of the human alpha-2 macroglobulin gene and its promotor.
A;Reference number: JN0262; MUID:92246939; PMID:1374237
A;Accession: JN0262
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-29 <MAT>
R;Sottrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Wierzbicki, D.M.; Jones, C.M.; Lon
J. Biol. Chem. 259, 8318-8327, 1984
A;Title: Primary structure of human alpha-2-macroglobulin. V. The complete structure.
A;Reference number: A92486; MUID:84239807; PMID:6203908
A;Accession: A92486
A;Molecule type: protein
A;Residues: 24-562,'E',564-974,'Z',976-999,'I',1001-1474 <SOT>

R;Sottrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Wierzbicki, D.M.; Jones, C.M.; Lonb
J. Biol. Chem. 260, 6500, 1985
A;Reference number: A92529
A;Contents: annotation; erratum
R;Virca, G.D.; Salvesen, G.S.; Travis, J.
Hoppe-Seyler's Z. Physiol. Chem. 364, 1297-1302, 1983
A;Title: Human neutrophil elastase and cathepsin G cleavage sites in the bait region of a
A;Reference number: A91713; MUID:84030513; PMID:6195065
A;Contents: annotation; inhibitory site
R;Sottrup-Jensen, L.; Lonblad, P.B.; Stepanik, T.M.; Petersen, T.E.; Magnusson, S.; Jornv
FEBS Lett. 127, 167-173, 1981
A;Title: Primary structure of the 'bait' region for proteinases in alpha-2-macroglobulin
A;Reference number: A91290; MUID:81212827; PMID:6165619
A;Contents: annotation; inhibitory site
R;Hall, P.K.; Nelles, L.P.; Travis, J.; Roberts, R.C.
Biochem. Biophys. Res. Commun. 100, 8-16, 1981
A;Title: Proteolytic cleavage sites on alpha-2-macroglobulin resulting in proteinase binc
A;Reference number: A90099; MUID:81255805; PMID:6167263
A;Contents: annotation; inhibitory site
R;Mortensen, S.B.; Sottrup-Jensen, L.; Hansen, H.F.; Petersen, T.E.; Magnusson, S.
FEBS Lett. 135, 295-300, 1981
A;Title: Primary and secondary cleavage sites in the bait region of alpha-2-macroglobulir
A;Reference number: A91299; MUID:82095610; PMID:6172288
A;Contents: annotation; inhibitory site
R;Dolmer, K.; Jenner, L.B.; Jacobsen, L.; Andersen, G.R.; Koch, T.J.; Thirup, S.; Sottrup
FEBS Lett. 372, 93-95, 1995
A;Title: Crystallisation and preliminary X-ray analysis of the receptor-binding domain of
A;Reference number: S66634; MUID:96032553; PMID:7556651
A;Accession: S66634
A;Molecule type: protein
A;Residues: 1337-1343 <DOL>
C;Comment: This inhibitor is able to inhibit all four classes (EC 3.4.21-3.4.24) of endop
s tertiary conformation, entrapping the proteinase. This results in hydrolysis of its thi
e and appears to unmask a receptor-specific binding site.
C;Comment: The entrapped enzyme remains active against low molecular weight substrates (a
r complex results in its rapid clearance from the circulation by receptor-mediated endocy
C;Comment: The wide specificity of this inhibitor is attributed to the primary sequence c
ities of a variety of plasma proteinases, form primary (residues 704-809) and secondary e
C;Genetics:
A;Gene: GDB:A2M
A;Cross-references: GDB:119639; OMIM:103950
A;Map position: 12p13.3-12p12.3
C;Complex: homotetramer; dimer of disulfide linked dimers
C;Superfamily: alpha-2-macroglobulin
C;Keywords: glycoprotein; plasma; proteinase inhibitor; thiolester bond
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-1474/Product: alpha-2-macroglobin #status experimental <MA2>
F;693-694,704-709,719-723,730-735/Region: inhibitory #status predicted
F;698-742/Region: bait region
F;48-86,251-299,269-287,278-431,595-771,642-689,821-849,847-883,921-1321,1079-1127,1352-1
F;55,70,247,396,410,869,991,1424/Binding site: carbohydrate (Asn) (covalent) #status expe
F;470,563/Disulfide bonds: interchain #status predicted
F;972-975/Cross-link: thiolester (Cys-Gln) #status experimental

```
  Query Match             19.6%;  Score 1440;  DB 1;  Length 1474;
  Best Local Similarity   29.1%;  Pred. No. 2.2e-73;
  Matches  451;  Conservative  266;  Mismatches  551;  Indels  284;  Gaps   59;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
              || :      :|:  |:::|  | ::          |  || :     ||| |  |       | ::
Db         15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy         72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
                :: ||||  |         :|     ::| :|::|       ::| | ||:   |     |  :   :
Db         74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy        127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
               :    ||:||||::|||  |  ||||:|::   :| |      : :: |:|||  |  |  ||  |  |  :
Db        125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy        186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
               |:    :| |||  |    | :  : ||        :|  :    |  |  |:||||||  :    |      ::
Db        185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243
```

```
Qy    244 KHLNGTITAKYTYGKPVKGDVTLTFLPL---------------------------SFWGK 276
Db    244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303
Qy    277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
Db    304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347
Qy    332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
Db    348 TKLS-FVKVDSHFRQG---IPFFG---------------QVRLVDGKGVPIP----NKVI 384
Qy    392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
Db    385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441
Qy    439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
Db    442 EEAHHTAY------------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489
Qy    493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
Db    490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549
Qy    541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
Db    550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608
Qy    601 ASNDITMENVVHEL-ELYNTGY--------------YLGMFMNS--------------FA 631
Db    609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 668
Qy    632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
Db    669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725
Qy    681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
Db    726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782
Qy    738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
Db    783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841
Qy    798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
Db    842 EQAPHCICANGRQ----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890
Qy    842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
Db    891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947
Qy    897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
Db    948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEVKSKAIGY 1007
Qy    957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
Db   1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067
Qy   1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
Db   1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127
Qy   1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
Db   1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186
Qy   1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
Db   1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245
Qy   1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
Db   1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304
Qy   1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
Db   1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353
Qy   1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
Db   1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413
Qy   1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
Db   1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465
```

RESULT 7
T43166
alpha-2-macroglobulin - Japanese lamprey
C;Species: Lampetra japonica (Japanese lamprey)
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 20-Jun-2000
C;Accession: T43166
R;Nonaka, M.
submitted to the EMBL Data Library, November 1992
A;Reference number: Z22323
A;Accession: T43166
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1503 <NON>
A;Cross-references: EMBL:D13567; PIDN:BAA02762.1
A;Experimental source: liver; clone Cl.3 and Cl1.10
C;Superfamily: alpha-2-macroglobulin

```
  Query Match             19.4%;  Score 1424;  DB 2;  Length 1503;
  Best Local Similarity   28.3%;  Pred. No. 1.8e-72;
  Matches  442;  Conservative  256;  Mismatches  600;  Indels  266;  Gaps   48;

Qy     10 AHLLCVCTAALAVAPGPR----FLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKT 65
Db      7 AFLLLLCALAAVVSADEQSSGHYLVFVPSELHALSSERLCVSLL-GVTGEVTFRATLHYK 65
Qy     66 ASNLTVSVLEAEGVFEKGSFKTLTLPS-----------LPLNSADEIYELRVTGRTQDEI 114
Db     66 DSRDTSHVAQNHTVVGEVDAGVLLQPSGGDLHHCFSFTVP-DVQGTTYANLVVRAAGEGL 124
Qy    115 LFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKP-YKTSLNILIKDPKS 173
Db    125 NFTKTHAVVVRKVKDVVFVQTDKPVYKPGQSVKFRVVTLDENFATVLKTYALIYIEDPQR 184
Qy    174 NLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQV---QVNDQTYYQSFQVSEYVLPKFE 230
Db    185 NRIAQWRNASGRAGIVQLELDMPSEPPLGTYNVMVVEQSGGDSVASHTFTVEEYVLPTFE 244
Qy    231 VTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSA 290
Db    245 VSIQTPSYLNYLDKSVTLKVCGRYTYGKPVHGAVNASVC----------------IQGQP 288
Qy    291 NFSFNDEEMKNVMD------SSNGLSEY-LDLSSPG--------PVEILTTVTESVTGIS 335
Db    289 RFWWREECIIPVCNEFFMKVGKDGCAEWQVDNAKPSNASCHTTHVLKVVAVLEEEGTGMK 348
Qy    336 RNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQ 395
Db    349 MKQKAEKNFETDITRISFVDMPSWYRHGLPIVGKVKVERPDGSPVP----HKLVSLIVKQ 404
Qy    396 RNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS---KS 452
Db    405 GNAAPPGSQHHTG-----ADGTFTFTI-DTGDFNRSDTIFLEATDPEFNSTAHPSVTYQQ 458
Qy    453 SMAVHSLFKSPSKTYIQLKTRDENIKVGS--PFELVVSGNKRLKE-------------- 495
Db    459 GYSTISAFFSPSDSFLQLDRVAHTLECGSSVPLRLLLV----LKENRSSAGDGHAGGVVH 514
```

```
Qy    496 --LSYMVVSRGQL---------VAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEII 543
            :: :|:|||  :          | :|| :|  | |  ::|: | | :: : : :: | :
Db    515 PVINVLVMSRGNILHTESFTMDNVYLGKAHSHHFQLKVKHSFAPTARLLAWMVFNN-ETV 573

Qy    544 SDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASN 603
          :| | :||    | ||: : :    :   |    :| |   : | |:  :  ||||| |:
Db    574 ADTLALPVHKCFPNKVSVGFEHEQELPGAMSTLSIRAS-PGSLCALRVVDKSVLLLKPEA 632

Qy    604 DITMENVVHELELYNTGYY-------------------LGMFMNSF-----AVFQECGL 638
           :: |::    |:: :  ||                   |::|: |        :|:  ||
Db    633 QLSAESIFQRLQVQDLTYYDYGIERWDCSPEHYRRRKRSYGVWMDDFQNSVHRLFKGMGL 692

Qy    639 WVLTDANLTKDYID------------GVYDNAEYAERFMEENEGHIVDIHDFSLGSS-- 683
          ||||  :  |  :|                 :    |       ||  :        :|      | :
Db    693 VVLTDTTV-KRSVDCHQLMHHRHRIAYSMETVASPGLRFSDRG-----NISPVGQGGAMF 746

Qy    684 ----PHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTP 739
                 ||::|||||||     :       ::  |  ||||||  | |:  |   |    | ||:
Db    747 KTVKETVREYFPETWIWDLYPVSESGLEQVAVKVPDSITEWKASAF-CSSPAGFGLSEVS 805

Qy    740 VELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEI 799
           |: | |||:   |||||:||| | | |:: |||   :::| :  |: | :
Db    806 -SLRVFTPFFVEPVLPYSVVRGETFPLAISVHNYLHSCLKIEVTLLDSEHFAV------- 857

Qy    800 NATGHQQ---TLLVPSEDGATVLFPIRPTHLGEIPITVTALS---------------PT 840
           | |  :    : : |   |     | | |  || : ::| | :
Db    858 -AAGGPRGAASCVCPDRKG-LFSFLIEPLALGTVNVSVRAAAVHSDEPCGNEVVVVPEEG 915

Qy    841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
           | | | : ::|: ||| |   | |·|  |       |        :      | | : ||  |
Db    916 AVDTVVRSVIVEPEGIPKELAYSSL--LCPKSPASERFVFNLPLPANVIVGSARAYATIA 973

Qy    901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
          ||::| ::     |: :| |||||||: ||||||| :||     |||  :::|||:|:| |
Db    974 GDIMGSALQNLDKLLTLPTGCGEQNMVKFAPNIYIQEYLQNSGQLTDAVRDKALNFLRVG 1033

Qy    961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
          |||:| |:|:| |:||||   |   |:|||:||||: |: |  :| : :: :  ::|| |
Db   1034 YQRQLTYKRDDHSYSAFGKSDDDGNTWLTAFVLKSFVRASKHIAVSEDHITGPFSWLVEH 1093

Qy   1021 QK-SNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNID---VQESIH--- 1073
          |  | | |  ||: :: ::||  |:|||||:  :||    : ||    |: ::
Db   1094 QNASTGCFISVGRLFNNAMKGGVSDDVSLTAYVTAALL-----ESNISGPVVERALGCLR 1148

Qy   1074 --FLESEFSRGISDNYTLALITYALS-SVGSPKAKEALNMLTWRAEQEGGMQFWVSSESK 1130
            |||     |: : ||| :|| | |   |   |:  : | || :|:|:  | :   |
Db   1149 PLVLES------SNAHLLALASYAFSLSADGATAQALRDALQSRAVTKGGLTHWQEANKK 1202

Qy   1131 LSDSWQPR--------------SLDIEVAAYALLS------HFLQFQTSEGIPIMRWLSR 1170
                 :                  |  :|  |||||                ||  | :::|||:
Db   1203 DDKDEEDEEGGFNFRRSYGGTTSAAVETTAYALLSRLAVPGQLAASATSSNIGVVQWLSK 1262

Qy   1171 QRNSLGGFASTQDTTVALKALSEFAAL-----------------MNTERTNI-QVTVTGP 1212
          |||: |||:||||| | |:||| ||||                  :::  | : :|::
Db   1263 QRNAYGGFSSTQDTVVGLQALSAFAALIDGDGGGGGSGGIKAVVLDSAHTLLREVSIDST 1322

Qy   1213 SS--------PSPLAVVQPMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAF 1264
          ::          | |:      :  : | | | | |: |:::  ||       |   :          |
Db   1323 NALILHQVQLPPPVPAPMVSSCTVEATGQGCALFQVSLKYNEPPKSSKPK---------F 1373

Qy   1265 DLDVAVKE---NKDDLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGFMVPSEAISLSET- 1319
           | |         :|           :||  |:  |   |  | |||:|| ::||:       ::       :|
Db   1374 SLSVETNPANCSKAARKGFQINVEVSYHGERGESNMALVEVKMISGYSAVKSSLKELQTF 1433

Qy   1320 ---VKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVR 1376
             ||||| |  :: :||| |::     :   :     | | | |:| : |||   |
Db   1434 YDFVKKVEVDASRVVIYLDKVDKKPIAIKLSVTQDIAVDNLQPATVRVYDYYATEDAATS 1493

Qy   1377 SYNS 1380
           |::
Db   1494 MYSA 1497
```

RESULT 8
S15904
alpha-1 proteinase inhibitor III, variant 1 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999
C;Accession: S15904; S52235; S29737; B34278; B29952; C29952; D29952; B27199
R;Regler, R.; Sickinger, S.; Schweizer, M.
FEBS Lett. 282, 368-372, 1991
A;Title: Differential regulation of the two mRNA species of the rodent negative acute pha
A;Reference number: S15904; MUID:91243832; PMID:1709877
A;Accession: S15904
A;Molecule type: mRNA
A;Residues: 1-1487 <REG>
A;Cross-references: EMBL:X52984; NID:g55561; PIDN:CAA37176.1; PID:g55562
R;Sickinger, S.; Schweizer, M.
submitted to the EMBL Data Library, May 1992
A;Reference number: S52235
A;Accession: S52235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-161 <SIC>
A;Cross-references: EMBL:X66454; NID:g609245; PIDN:CAA47070.1; PID:g609246
R;Thomsen, N.K.; Sottrup-Jensen, L.
Arch. Biochem. Biophys. 300, 327-334, 1993
A;Title: alpha-macroglobulin domain structure studied by specific limited proteolysis.
A;Reference number: S29737; MUID:93143334; PMID:7678727
A;Accession: S29737
A;Molecule type: protein
A;Residues: 25-32,276-280,'XG',314-320,'X',322,'SL',717,'VVAV',722-726,'Q',728,1205-1218
R;Enghild, J.J.; Salvesen, G.; Thogersen, I.B.; Pizzo, S.V.
J. Biol. Chem. 264, 11428-11435, 1989
A;Title: Proteinase binding and inhibition by the monomeric alpha-macroglobulin rat alpha
A;Reference number: A34278; MUID:89291900; PMID:2472396
A;Accession: B34278
A;Molecule type: protein
A;Residues: 684-715,'N',717-733 <ENG>
R;Braciak, T.A.; Northemann, W.; Hudson, G.O.; Shiels, B.R.; Gehring, M.R.; Fey, G.H.
J. Biol. Chem. 263, 3999-4012, 1988
A;Title: Sequence and acute phase regulation of rat alpha-1-inhibitor III messenger RNA.
A;Reference number: A92709; MUID:88153707; PMID:2831216
A;Accession: B29952
A;Molecule type: mRNA
A;Residues: 1-3,'D',5-86,'V',88-136,'M',138-169,361-364,'M',366,'V',368-397,'H',399-727,'
A;Cross-references: GB:J03552; GB:M22359
A;Note: the authors translated the codon ATG for residue 174 as Ile
A;Accession: C29952
A;Molecule type: mRNA
A;Residues: 80-262,269-361 <BR2>
A;Cross-references: GB:M22360; NID:g202583; PIDN:AAA40633.1; PID:g554402
A;Accession: D29952
A;Molecule type: mRNA
A;Residues: 315-409;1064-1079,'LAASGAPDHC',1090,'TMP' <BR3>
A;Cross-references: GB:M22361; GB:M22362
R;Schweizer, M.; Takabayashi, K.; Geiger, T.; Laux, T.; Biermann, G.; Buhler, J.M.; Gauth
Eur. J. Biochem. 164, 375-381, 1987
A;Title: Identification and sequencing of cDNA clones for the rodent negative acute-phase
A;Reference number: A27199; MUID:87190405; PMID:2436907
A;Accession: B27199
A;Molecule type: mRNA
A;Residues: 1416-1483,'N',1485-1487 <SCH>
A;Cross-references: GB:M28297
A;Experimental source: liver, clone pI3cDNA3
C;Genetics:
A;Introns: 29/2; 90/3; 144/1
C;Superfamily: alpha-2-macroglobulin
C;Keywords: proteinase inhibitor; thiolester bond
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-1487/Product: alpha-1 inhibitor III variant 1 #status predicted <MAT>
F;648-723/Region: bait region #status predicted
F;985-988/Cross-link: thiolester (Cys-Gln) #status predicted

Query Match 19.3%; Score 1417; DB 2; Length 1487;

```
  Best Local Similarity   29.0%;  Pred. No. 4.5e-72;
  Matches  454;  Conservative  244;  Mismatches  576;  Indels  290;  Gaps   55;

Qy         13 LCVCTAALAVAPGPRFLVTAPGIIRPGGN----VTIGVELLEHCPSQVTVKAELLKTASN 68
              ||: :| ||  |      |              ||      |  :  :|   | :: :
Db          9 LCLFSALLAFLPPASLL---------NGNSKYMVLVPSQLYTETPEKICLHLYHLNETVT 59

Qy         69 LTVSVLEAEG--------VFEKGSFKTL--TLPSLPLNSADEIYELRVTGRTQDEILFSN 118
              :| |::    |         |     |    |  |:| || :  :|  :: :|      ||
Db         60 VTASLISQRGTRKLFDELVVDKDLFHCLSFTIPRLPSSEEEESLDINIEGAKHK---FSE 116

Qy        119 STRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-----IKDPKS 173
                :  : |   ||:||||  :|||  | ||||:|::   :   |       || |      |:|||
Db        117 RRVVLVKNKESVVFVQTDKPVYKPGQSVKFRVVSMDKNLHP----LNELFPLAYIEDPKM 172

Qy        174 NLIQQWLSQQSDLGVISKTFQLSSHPILGDWSI----QVQVNDQTYYQSFQVSEYVLPKF 229
               | | ||    ::: |:    :| ||: || |  : |       |  | ::      || |  |:|||:|
Db        173 NRIMQWQDIKTENGLKQLSFSLSAEPIQGPYKIVILKQSGVKEE---HSFTVMEFVLPRF 229

Qy        230 EVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFL---PLSFWGKKKNITKTFKI 286
               | :: |     |:  : :|  |   | ||||||| | |  ::        | ||  : |:  |
Db        230 GVDVKVPNAISVYDEIINVTACAIYTYGKPVPGHVKISLCHGNP-SFSSETKSACKE--- 285

Qy        287 NGSANFSFNDEEMKNVMDSSNGLS-------------EYLDLSSPGPVEILTTVTESVTG 333
                        |  |: |        || |                    ||  :   ||||  ||
Db        286 --------EDSELDN-----NGCSTQEVNITEFQLKENYLKMHQ--AFHVNATVTEEGTG 330

Qy        334 ISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITV 393
               :  |    ::            |        :   :   :   |      :::     |: :  |:     |  |
Db        331 SEFSGSGRIEVERTRNKFLFLKADSHFRHGIPFFVKIRLVDIKGDPIPNEQ----VFIKA 386

Qy        394 TQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSS 453
               :  ||   :     |  |                       | |:   :    |  | :| |     : |
Db        387 QEAGYTNATTTDQHGLAK----------------FSIDTSSISGYS-LNIKVYH--KEES 427

Qy        454 MAVHS---------------LFKSPSKTYIQLKTR------DENIKVGSPFEL--VVSGN 490
                 :||                    | ||:|| | |          ::    |  :  | |    |  |
Db        428 SCIHSSCTAERHAEEHHTAYAVYSLSKSYIYLDTEAGVLPCNQIHTVQAHFILKGQVLGV 487

Qy        491 KRLKELSYMVVSRGQLVAVGK------------QNSTMFSLTPENSWTPKACVIVYYIED 538
                |:|:::| ::  |                 |  :     :    |    |   |  :::| |
Db        488 LPQIVFHYLVMAQGSILQTGNHTHQVEPGVSQVQGNFALEIPVEFSMVPVAKMLIYTILP 547

Qy        539 DGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNL 598
              |||:|:|  :     |:    :  :||: | :|   :|   :| :|:  :  |  |:|  :|  |||:|| |
Db        548 DGEVIADSVTFQVEKCLRNKVHLSFSPSQSLPASQTHMRVTAS-PQSLCGLRAVDQSVLL 606

Qy        599 MNASNDITMENVVHEL-------------------------------EL-YNTGYYLGMF 626
              :     :::  :::::|                                || |:    |    |
Db        607 LKPEAELS-PSLIYDLPGMQDSNFIPSSYHPFEDEYDCLMYQPRDTEELTYSVPY--GRE 663

Qy        627 MNSFAVFQECGLWVLTDANL---TKDY-IDGVYDNAEYAERFMEENEGHIVDIHDFSLGS 682
               : :   :: ||    |: :     |   | :: |    :|     |      |      ::     :
Db        664 KDVYRYVRDMGLTAFTNLKIKHPTYCYEMNMVVLSAPAVESELSPRGGEF-EMMPLGVNK 722

Qy        683 SP------------------HVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATG 724
              ||                     ·:| :||||||      :           | |:||||:|| | |
Db        723 SPLPKEPPRKDPPPKDPVIETIRNYFPETWIWDLVTVNSSGVTEVEMTVPDTITEWKAGA 782

Qy        725 FVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVII 784
               :| | ||||::          |||||||: | :|||||||| |: |: |||   : : | :
Db        783 LCLSNDTGLGLSSV-ATLQAFQPFFVELTMPYSVIRGEAFMLKATVMNYLPTSLPMAVQL 841

Qy        785 EKSDKFDILMTSSE-----INATG-HQQTLLVPSEDGATVLFPI------RPTHLGEIPI 832
               | |  |    :    ::         : | |  |   : ||   :       | |  :          |
Db        842 EASPDFTAVPVGNDQDSYCLGANGRHTSSWLVTPKSLGNVNFSVSVEAQQSPELCGSQVA 901

Qy        833 TVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGS 892
               ||           | | ::::|: |||:| :: | ||  :|  |      :|||    ||  |   |
Db        902 TVPETG--RKDTVVKVLIVEPEGIKKEHTFSSLLCASDAELS---ETLSLLLPPTVVKDS 956

Qy        893 ERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEK 952
                 |    : :||:|  :|      :||:||||||||||: ||||||:| ||  : :|||: :| |
Db        957 ARAHFSVMGDILSSAIKNTQNLIQMPYGCGEQNMVLFAPNIYVLKYLNETQQLTEKIKSK 1016

Qy        953 ALSFMRQGYQRELLYQREDGSFSAFGNYD--PSGSTWLSAFVLRCFLEADPYIDIDQNVL 1010
              || ::| ||||||  |: :|||:||||:::       |:|||:||||: | :|   :| ||:: :
Db       1017 ALGYLRAGYQRELNYKHKDGSYSAFGDHNGQGQGNTWLTAFVLKSFAQARAFIFIDESHI 1076

Qy       1011 HRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNID--V 1068
                 :|||     || :|  |     |  : :: ::||       :||:|||    :||               |:  |   |
Db       1077 TDAFTWLSKQQKDSGCFRSSGSLFNNAMKGGVDDEITLSAYITMALL--ESSLPDTDPVV 1134

Qy       1069 QESIHFLES-----EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQ 1122
               :::  ||:        |   |    |   ||   |:  |  :    |:         |   |    |    |  :|
Db       1135 SKALGCLEASWETIEQGRNGSFVYTKTLMAYAFALAGNQEKRNEILKSLDKEAIREDNSI 1194

Qy       1123 FW------VSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIP----------IMR 1166
               |             ||   |       |   | ::|::|| :|:         |::   |                 ::
Db       1195 HWERPQKPTKSEGYLYTP-QASSAEVEMSAYVVLARL----TAQPAPSPEDLALSMGTIK 1249

Qy       1167 WLSRQRNSLGGFASTQDTTVALKALSEF-AALMNTERTNIQVTVTGPSSPSPLAVVQ--- 1222
              ||::|:|| |||:|||||| ||| |||:: ||   :    :         ||:         |  |     |
Db       1250 WLTKQQNSHGGFSSTQDTVVALDALSKYGAATFSKSQKTPLVTIQSSGSFSQKFQVDNSN 1309

Qy       1223 ------------PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAV 1270
                          |      :| :| |       |  :  ||:                :  | || |  |
Db       1310 RLLLQQVSLPDIPGNYTVSVSGEGCVYAQTTLRYNMPLE---------KQQPAFALKVQT 1360

Qy       1271 K----ENKDDLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGFMVPSEAISLSETVKKVE- 1324
               |       |   :::   |::|     | |  : :| :||||           | | · ||||:|
Db       1361 VPLTCNNPKGQNSFQISLEISYTGSRPASNMVIADVKMLSGF------IPLKPTVKKLER 1414

Qy       1325 YDH---------GKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVR 1376
                  :|                    :  ||||  |            :  ::     |   |  |  |  |  |  :  ||||       |
Db       1415 LEHVSRTEVTTNNVLLYLDQVTNQTLSFSFIIQQDIPVKNLQPAIVKVYDYYETDEVAFA 1474

Qy       1377 SYNS 1380
                |:|
Db       1475 EYSS 1478
```

RESULT 9
A26122
alpha-2-macroglobulin precursor [validated] - rat
N;Alternate names: alpha-2M
C;Species: Rattus norvegicus (Norway rat)
C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 31-Dec-2000
C;Accession: A26122; A05278; S03431; S33843; A22614; I52403; I54013
R;Gehring, M.R.; Shiels, B.R.; Northemann, W.; de Bruijn, M.H.L.; Kan, C.C.; Chain, A.C.
J. Biol. Chem. 262, 446-454, 1987
A;Title: Sequence of rat liver alpha-2-macroglobulin and acute phase control of its mess
A;Reference number: A26122; MUID:87083488; PMID:2432068
A;Accession: A26122
A;Molecule type: mRNA
A;Residues: 1-1472 <GEH>
A;Cross-references: GB:J02635; NID:g202591; PIDN:AAA40636.1; PID:g202592
R;Hayashida, K.; Okubo, H.; Noguchi, M.; Yoshida, H.; Kangawa, K.; Matsuo, H.; Sakaki, Y
J. Biol. Chem. 260, 14224-14229, 1985
A;Title: Molecular cloning of DNA complementary to rat alpha-2--macroglobulin mRNA.
A;Reference number: A05278; MUID:86033908; PMID:2414291
A;Accession: A05278
A;Molecule type: mRNA
A;Residues: 178-227;420-489,'L',491-526 <HAY>
R;Kunz, D.; Zimmermann, R.; Heisig, M.; Heinrich, P.C.
Nucleic Acids Res. 17, 1121-1138, 1989
A;Title: Identification of the promoter sequences involved in the interleukin-6 dependen
A;Reference number: S03431; MUID:89160243; PMID:2466233
A;Accession: S03431
A;Molecule type: DNA
A;Residues: 1-22,'V',24-164 <KUN>
A;Cross-references: EMBL:X13983
A;Note: the authors translated the codon GTG for residue 23 as Leu
R;Warburton, M.J.; Coles, B.; Dundas, S.R.; Gusterson, B.A.; O'Hare, M.J.
Eur. J. Biochem. 214, 803-809, 1993
A;Title: Hydrocortisone induces the synthesis of alpha(2)-macroglobulin by rat mammary m

```
A;Reference number: S33843; MUID:93307297; PMID:7686489
A;Accession: S33843
A;Molecule type: protein
A;Residues: 28-40,'R',42-48 <WAR>
R;Northemann, W.; Heisig, M.; Kunz, D.; Heinrich, P.C.
J. Biol. Chem. 260, 6200-6205, 1985
A;Title: Molecular cloning of cDNA sequences for rat alpha-2-macroglobulin and measureme
A;Reference number: A22614; MUID:85207604; PMID:2581948
A;Accession: A22614
A;Status: significant sequence differences
A;Molecule type: mRNA
A;Note: the authors translated the codon ATT for residue 38 as Thr
R;Northemann, W.; Shiels, B.R.; Braciak, T.A.; Hanson, R.W.; Heinrich, P.C.; Fey, G.H.
Biochemistry 27, 9194-9203, 1988
A;Title: Structure and acute-phase regulation of the rat alpha 2-macroglobulin gene.
A;Reference number: I52403; MUID:89207509; PMID:2468362
A;Accession: I52403
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-146 <RES>
A;Cross-references: GB:M23567; NID:g202598; PIDN:AAA77658.1; PID:g554404
R;Tsuchiya, Y.; Hattori, M.; Hayashida, K.; Ishibashi, H.; Okubo, H.; Sakaki, Y.
Gene 57, 73-80, 1987
A;Title: Sequence analysis of the putative regulatory region of rat alpha-2-macroglobuli
A;Reference number: I54013; MUID:88112864; PMID:2448189
A;Accession: I54013
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-23,'A',25-227 <RE2>
A;Cross-references: GB:M22670; NID:g205388; PIDN:AAA41592.1; PID:g205390
C;Genetics:
A;Introns: 33/2; 94/3; 147/1; 164/3; 171/3
C;Superfamily: alpha-2-macroglobulin
C;Keywords: glycoprotein; plasma; proteinase inhibitor; thiolester bond
F;1-27/Domain: signal sequence #status predicted <SIG>
F;28-1472/Product: alpha-2-macroglobulin #status experimental <MAT>

  Query Match            19.2%;  Score 1412.5;  DB 2;  Length 1472;
  Best Local Similarity  27.6%;  Pred. No. 7.9e-72;
  Matches  426;  Conservative  275;  Mismatches  580;  Indels  265;  Gaps   51;

Qy          5 PLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLK 64
              |||   || : | | |     | ::| | ::  |        |   |   | |: |
Db         13 PLLLRLLLLLLPTDASA-PQKPIYMVMVPSLLH-AGTPEKACFLFSHLNETVAVRVSLES 70

Qy         65 TASN--LTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEI--YELRVTGRTQDEILFSNST 120
              |   |    :: : :|   ||      ::| :|:|:  : ::| | |  |     :
Db         71 VHGNQSLFTDLVVDKDLFHCTSF------TVPQSSSDELMFFTVQVKGATHE---FRRQS 121

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQW 179
               :  :  |     || || |||: :||  | |:||:|:|  | |   : :| |:|||:| | ||
Db        122 TVLVKKKESLVFAQTDKPIYKPGQTVRFRVVSLDESFHPLNELIPLLYIQDPKNNRIAQW 181

Qy        180 LSQQSDLGVISKTFQLSSHPILGDWSIQVQV-NDQTYYQSFQVSEYVLPKFEVTLQTPLY 238
                :  :|:   :| |||  |   |  : : :: : :|    |  | |:||||||| :  |
Db        182 QNFNLEGGLKQLSFPLSSEPTQGSYKVVIRTESGRTVEHPFSVEEFVLPKFEVRVTVPET 241

Qy        239 CSMNSKHLNGTITAKYTYGKPVKGDVTLTFL-----PLSFWGKKKNI------------ 280
               ::  : :| ::  ||||||  | ||:         | :|::
Db        242 ITILEEEMNVSVCGIYTYGKPVPGRVTVNICRKYSNPSNCFGEESVAFCEKLSQQLDGRG 301

Qy        281 -------TKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTG 333
                     ||:|::     :     :   :  |:  |     :  |      || :| ::|
Db        302 CFSQLVKTKSFQLK-RQEYEMQLDVHAKIQEEGTGVEE----TGKG----LTKITRTITK 352

Qy        334 ISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITV 393
              :|                :  |    :   :   :   |    |  :   |    :
Db        353 LS----------------FVNVDSHFRQGIPFVGQVLLVDGRGTPIP------------ 383

Qy        394 TQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILED---SSELQLKAYF--- 447
                   |   : |::           ||   || |  :  :  | | |     : |::| :
Db        384 ----YETIFIGAD------EANLYINTTTDKHGLAR--FSINTDDIMGTSLTVRAKYKDS 431

Qy        448 --------LGSKSSMAVHSLFK--SPSKTYIQLKTRDENIKVGSPFEL----VVSGN--K 491
                      | ::   | |:  :     |||::::| |::  : ::       |:     :::|  :
Db        432 NACYGFRWLTEENVEAWHTAYAVFSPSRSFLHLESLPDKLRCDQTLEVQAHYILNGEAMQ 491

Qy        492 RLKEL--SYMVVSRGQLVAVG------KQNSTM--FS--LTPENSWTPKACVIVYYIEDD 539
               ||||   |:::::| :|  |         ||         ||  :: |      | | :::| |  :
Db        492 ELKELVFYYLMMAKGGIVRAGTHVLPLKQGQMRGHFSILISMETDLAPVARLVLYAILPN 551

Qy        540 GEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLM 599
              ||::  |    |  ::     ||: | :         |: :   | :  :   | |: |: |||:|| ||
Db        552 GEVVGDTAKYEIENCLANKVDLVFRPNSGLPATRALLSV-MASPQSLCGLRAVDQSVLLM 610

Qy        600 NASNDITMENV--------------------------VHELELYNTGYYLGMFMNS---- 629
                :::   :                            | : :  |   |      |:
Db        611 KPETELSASLIYDLLPVKDLTGFPQGADQREEDTNGCVKQNDTYINGILYSPVQNTNEED 670

Qy        630 -FAVFQECGLWVLTDANLTKDYI-DGVYDNAEYAERFMEENEGHIVDIHDFSLGSSPHVR 687
                :   ::  || |  | ::|:  |   :  :  :  ||        :   ::  |:    :  |    :    |
Db        671 MYGFLKDMGLKVFTNSNIRKPKVCERLRDNKGIPAAYHLVSQSHMDAFLESSESPTETRR 730

Qy        688 KHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQP 747
               :||||||        :        | ||||||:|| | |    | :| |  ||||:    |: ||||
Db        731 SYFPETWIWDLVVVDSAGVAEVEVTVPDTITEWKAGAFCLSNDTGLGLSPV-VQFQAFQP 789

Qy        748 FFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQT 807
              ||: |  :|||||||| |  |: :|: |||       | | :|  |      ||         |       |
Db        790 FFVELTMPYSVIRGEAFTLKATVLNYLPTCIRVAVQLEASP--DFLAAPEEKEQRSH--- 844

Qy        808 LLVPSEDGATVLFPIRPTHLGEIPITVTALSPTA---------------SDAVTQMILV 851
                :          |  :  : |  ||  :  ||:|  :  :                     | : : :||
Db        845 -CICMNQRHTASWAVIPKSLGNVNFTVSAEALNSKELCGNEVPVVPEQGKKDTIIKSLLV 903

Qy        852 KAEGIEKSYSQSILLDLTDNRLQSTL-----KTLSFSFPPNTVTGSERVQITAIGDVLGP 906
               : |:|              :: |:| |    :        : ::  |  |  |  |  |   :  |  :|:|:|
Db        904 EPEGLEN--------EVTFNSLLCPMGAEVSELIALKLPSDVVEESARASVTVLGDILGS 955

Qy        907 SINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELL 966
              ::     |::||||||||||: ||||||:|||| :  :||  |  : :  |  ||:::  |||:|
Db        956 AMQNTQDLLKMPYGCGEQNMVLFAPNIYVLDYLNETQQLTQEIKTKAIAYLNTGYQRQLN 1015

Qy        967 YQREDGSFSAFGNYDPSG----STWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQK 1022
              |:  |||:|||| |  |       :|||:||||: | :|  || ||:   :  :    ||   ||
Db       1016 YKHRDGSYSAFG--DKPGRNHANTWLTAFVLKSFAQARKYIFIDEVHITQALLWLSQQQK 1073

Qy       1023 SNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEF--- 1079
               ||| |    | :::: ::|| :   |||:|||  :||        :  |: :: |:: :
Db       1074 DNGCFRSSGSLLNNAMKGGVEDEVTLSAYITIALLEMSLPVTHPVVRNALFCLDTAWKSA 1133

Qy       1080 --SRGISDNYTLALITYALSSVGSPKA-KEALNMLTWRAEQEGGMQFWVSSESKLSDS-- 1134
                | |  || ||: ||  : |:        ||  |   |  |   |:|     :  ||  |
Db       1134 RGGAGGSHVYTKALLAYAFALAGNQDTKKEILKSLDEEAVKEEDSVHWTRPQ-KPSVSVG 1192

Qy       1135 --WQPR--SLDIEVAAYALLSHFLQFQTSEGIP----------IMRWLSRQRNSLGGFAS 1180
                :||:  |::|:| :: ||  ||::        :   :           |::|:|:::|:|  ||:|:
Db       1193 LWYQPQATSAEVEMTAYVLLAYL----TTEPAPTQEDLTAAMLIVKWLTKQQNSHGGFSS 1248

Qy       1181 TQDTTVALKALSEF-AALMNTERTNIQVTVTGPSSPSPLAVVQ--------------PM 1224
              |||| ||| |||::: :|       :    |||:      :  |   |
Db       1249 TQDTVVALHALSKYGSATFTRAKKAAQVTIRSSGTFSTKFQVNNNNQLLLQRVTLPTVPG 1308

Qy       1225 AVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK----ENKDDLNHV 1280
                :   |  |    | :: |:|            : :|  |   ||  |:        :|  |
Db       1309 DYTVKVTGEGCVYLQTSLKYSV-----------LPREEEFPFAVVVQTLPGTCEDPKAHT 1357

Qy       1281 DLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISLSET---VKKVEYDHGKLNLYL 1334
                :  :  |  |    |  ||: :| ::|||:          : :  |    |  :  |   :  :  :||
Db       1358 SFQISLNISYTGSRSESNMAIADVKMVSGFIPLKPTVKMLERSVHVSRTEVSNNHVLIYL 1417

Qy       1335 DSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
              | |:       ::   ::   :  :  : | : |  ||||      ||  |::
Db       1418 DKVSNQTVNLSFTVQQDIPIRDLKPAVVKVYDYYEKDEFAVAKYSA 1463
```

```
RESULT 10
B41185
alpha-2 macroglobulin homolog MUG2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 16-Jul-1999
C;Accession: B41185
R;Overbergh, L.; Torrekens, S.; Van Leuven, F.; Van den Berghe, H.
J. Biol. Chem. 266, 16903-16910, 1991
A;Title: Molecular characterization of the murinoglobulins.
A;Reference number: A41185; MUID:91358495; PMID:1840592
A;Accession: B41185
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1451 <OVE>
A;Cross-references: GB:M65238; NID:g199888; PIDN:AAA73041.1; PID:g199889
C;Superfamily: alpha-2-macroglobulin

  Query Match             19.1%;  Score 1402;  DB 2;  Length 1451;
  Best Local Similarity   27.7%;  Pred. No. 3.1e-71;
  Matches  439;  Conservative  250;  Mismatches  527;  Indels  370;  Gaps   54;

Qy         13 LCVCTAALAVAPG-------PRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKT 65
Db          9 LCLFSVLLAFLPSASSLNGDSKYMVLVPS-----------QLYTETPEKICLHLYHLNE 56

Qy         66 ASNLTVSVLEAEG--------VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEI 114
Db         57 TVTVTASLVSQTGRRNLFDELVVDKDLFQCVSFIIPTLNSPDEEEFLY-VDIKGPTHE-- 113

Qy        115 LFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNI-LIKDPKS 173
Db        114 -FSKRNAVLVKNKESVVFVQTDKPVYKPGQSVKFRVVSMDKTLRPLNELLPLAYIEDPKK 172

Qy        174 NLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQ----VNDQTYYQSFQVSEYVLPKF 229
Db        173 NRIMQWRDIKTENGLKQMSFSLAAEPIQGPYKIVVHKQSGVKEE---HSFTVMEFVLPRF 229

Qy        230 EVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGS 289
Db        230 NVDLKVPNAISVNDEVLQVTVCGKYTYGKPVPGQVKISICHETEAGCKE-VNSKLDNNGC 288

Qy        290 ANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRN---------VST 340
Db        289 STQEVNITELQ-----SKKRNYEVQL-----FHVNATVTEEGTGLEFNGYGTTKIERITN 338

Qy        341 NVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTE 400
Db        339 KLIFLKAD---------SHFRHGIPFFVKVRLVDIKGDPIPNE----RVFIKAQVLGYTS 385

Qy        401 YWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKS-------- 452
Db        386 TTTTDQHGLAK--------FSIDTAGF---------SGSSLHIKVNHKGKDSCYFFYCME 428

Qy        453 ---------SMAVHSLFKSPSKTYIQLKTRDENI-------KVGSPFELVVSGNKR-LKE 495
Db        429 ERYASAEHVAYAVYSL----SKSYIYLVKETSSILPCNQIHTVQAHF--ILKGDLGVLKE 482

Qy        496 L--SYMVVSRGQLVAVGKQNSTM------------FSLTPENSWTPKACVIVYYIEDDGE 541
Db        483 LVFYYLVMAQGSIIQTGNHTHQVEPGEAPVKGNFDLEIPVEFSMAPMAKMLIYTILPDGE 542

Qy        542 IISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNA 601
Db        543 VIADSVNFEIEKCLRNKVDLSFSSSQSLPASQTRLQVTAS-PQSLCGLRAVDQSVLLLKP 601

Qy        602 SNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYA 661
Db        602 EDELSPS------WIYNLP---GMQHNKF----------IPSSSLSEDREDCILYSSWVA 642

Qy        662 E---------------RFMEENEGHIVDIHDFS-------------------------- 679
Db        643 EKHTDWVPHGREKDVYRYVED-----MDLKAFTNLKIKLPKICFDSAPMSGPRGKFDLAF 697

Qy        680 ---------LGSSPH--------------------VRKHFPETWIWLDTNMGYRIYQEFE 710
Db        698 SSEVSGTLQKGSSKRPQPEEPPREDPPPKDPLAETIRKYFPETWVWDIVTVNSTGVAEVE 757

Qy        711 VTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITI 770
Db        758 MTVPDTITEWKAGALCLSNDTGLGLSSV-VPLQAFQPFFVEVSLPYSVVRGEAFMLKATV 816

Qy        771 FNYLKDATEVKVIIEKSDKFDILMTSSE-----INATG-HQQTLLVPSEDGATVLFPIRP 824
Db        817 MNYLPTSMRMSVQLEASPDFTAVPVGDDHDSYCLSANGRHTSSWLV------------TP 864

Qy        825 THLGEIPITVTALSPTAS----------------DAVTQMILVKAEGIEKSYSQSILLDL 868
Db        865 KSLGNVNFSVSVEAQQSSEPCGSEVATVPATGRKDTVVKVLIVEPEGIKQEHTFNSLFCA 924

Qy        869 TDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMIN 928
Db        925 SDAEIS---EKMSLVLPPTVVKDSARAHFSVMGDILSSAIKNTQNLLHMPYGCGEQNMVL 981

Qy        929 FAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFG--NYDPSGST 986
Db        982 FAPNIYVLKYLDKTQQLTQKIKTKALGFLRAGYQRELNYKHKDGSYSAFGDQNGEREGNT 1041

Qy        987 WLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPV 1046
Db       1042 WLTAFVLKSFAQARAFIFIDESHITHAFTWLSQQQKDNGCFRSSGSLFHNDI----KHPV 1097

Qy       1047 TLTAYIVTSLLGYRKYQPNIDVQESIHFLES-----EFSRGISDNYTLALITYALSSVGS 1101
Db       1098 ---------------------VSKALSCLESSWKTIEQGRNANFVYTKALMAYAFALAGN 1136

Qy       1102 -PKAKEALNMLTWRAEQEGGMQFWV------SSESKLSDSWQPRSLDIEVAAYALLSHFL 1154
Db       1137 QDKRNEILKSLDEEAIKEDNSIHWERPQKPRKSEHNLYKP-QASSIEVEMNAYVVLARL- 1194

Qy       1155 QFQTSEGIP----------IMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN 1204
Db       1195 ---TAQPAPSPEDLTLSRSTIMWLTKQQNSNGGFSSTQDTVVALDALSKYGAVTFSRRQK 1251

Qy       1205 IQ-VTVTGPSSPSPLAVVQ---------------PMAVNISANGFGFAICQLNVVYNVKA 1248
Db       1252 TSLVTIQSTGSFSQKFQVENSNCLLLQQVPLPDIPGDYTISVSGEGCVYAQTTLRYNMHL 1311

Qy       1249 SGSSRRRRSIQNQEAFDLDVAVK----ENKDDLNHVDLNVCTSFSGP-GRSGMALMEVNL 1303
Db       1312 E---------KQQSAFALRVQTVPLTCNNPKGHNSFQISLEISYTGSRPASNMVIADVKM 1362

Qy       1304 LSGFMVPSEAISLSETVKKV----------EYDHGKLNLYLDSVNETQFCVNIPAVRNFKV 1354
Db       1363 LSGF------IPLKPTVKKLERLEHISRTEVSNNNVLLYLDQVTNQTLAFSFIIQQDISV 1416

Qy       1355 SNTQDASVSIVDYYEPRRQAVRSYNS 1380
Db       1417 RNLQPAIVKVYDYYETDEVAYAEYSS 1442

RESULT 11
S13495
pregnancy zone protein - human
C;Species: Homo sapiens (man)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 16-Jul-1999
C;Accession: S13495; S09106; A34230; PS0077; S21637
R;Devriendt, K.; van den Berghe, H.; Cassiman, J.J.; Marynen, P.
Biochim. Biophys. Acta 1088, 95-103, 1991
A;Title: Primary structure of pregnancy zone protein. Molecular cloning of a full-length
A;Reference number: S13495; MUID:91113734; PMID:1989698
A;Accession: S13495
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-1482 <DEV>
A;Cross-references: EMBL:X54380; NID:g35824; PIDN:CAA38255.1; PID:g35825
R;Marynen, P.; Devriendt, K.; van den Berghe, H.; Cassiman, J.J.
```

FEBS Lett. 262, 349-352, 1990
A;Title: A genetic polymorphism in a functional domain of human pregnancy zone protein:
n.
A;Reference number: S09106; MUID:90242963; PMID:1692292
A;Accession: S09106
A;Molecule type: DNA
A;Residues: 668-690,'M',692-753 <MAR>
A;Cross-references: EMBL:X51541
R;Sottrup-Jensen, L.; Sand, O.; Kristensen, L.; Fey, G.H.
J. Biol. Chem. 264, 15781-15789, 1989
A;Title: The alpha-macroglobulin bait region. Sequence diversity and localization of cle
A;Reference number: A34230; MUID:89380162; PMID:2476433
A;Accession: A34230
A;Molecule type: protein
A;Residues: 670-752,'Q',754-759 <SOT>
C;Genetics:
A;Gene: GDB:PZP
A;Cross-references: GDB:120330; OMIM:176420
A;Map position: 12p13-12p12.2
A;Introns: 73/3; 116/2; 146/3; 169/3; 703/1; 753/2
A;Note: the list of introns may be incomplete
C;Superfamily: alpha-2-macroglobulin
F;685-735/Region: bait region

```
  Query Match             19.0%;  Score 1396.5;  DB 2;  Length 1482;
  Best Local Similarity   28.3%;  Pred. No. 6.5e-71;
  Matches  432;  Conservative  280;  Mismatches  587;  Indels  225;  Gaps   53;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
Db         14 LLILLSASDSNSTEPQYMVLVPSLLHTEAPKK-GCVLLSHLNETVTVSASLESGRENRS- 71

Qy         72 SVLEAEGVFEKGSFK--TLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRI 129
Db         72 --LFTDLVAEKDLFHCVSFTLPRISASSEVAFLSIQIKGPTQD--FRKRNTVLVLNTQSL 127

Qy        130 SVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSDLGV 188
Db        128 -VFVQTDKPMYKPGQTVRFRVVSVDENFRPRNELIPLIYLENPRRNRIAQWQSLKLEAGI 186

Qy        189 ISKTFQLSSHPILGDWSIQVQVNDQTYYQ-SFQVSEYVLPKFEVTLQTPLYCSMNSKHLN 247
Db        187 NQLSFPLSSEPIQGSYRVVVQTESGGRIQHPFTVEEFVLPKFEVKVQVPKIISIMDEKVN 246

Qy        248 GTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSN 307
Db        247 ITVCGEYTYGKPVPGLAT-----VSLCRKLSRVLNCDKQEVCEEFS---QQL-----NSN 293

Qy        308 GL------SEYLDLSSPG---PVEILTTVTESVTGI---SRNVS--TNVFFKQHDYIIEF 353
Db        294 GCITQQVHTKMLQITNTGFEMKLRVEARIREEGTDLEVTANRISEITNIVSK-----LKF 348

Qy        354 FDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSG----- 408
Db        349 VKVDSHFRQGIPFFA--QVLLVDGKGVPIPNKL--FFISVNDANYYSNATTNEQGLAQFS 404

Qy        409 -NQKMEAVQKI---NYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPS 464
Db        405 INTTSISVNKLFVRVFTVHPNLCFHYSW-VAEDHQGAQ--------HTANRVFSL----S 451

Qy        465 KTYIQLKTRDENIKVG------SPFELVVSGNKRLKELS--YMVVSRGQLVAVG------ 510
Db        452 GSYIHLEPVAGTLPCGHTETITAHYTLNRQAMGELSELSFHYLIMAKGVIVRSGTHTLPV 511

Qy        511 ----KQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV 566
Db        512 ESGDMKGSFALSFPVESDVAPIARMFIFAILPDGEVVGDSEKFEIENCLANKVDLSFSPA 571

Qy        567 KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELE----------- 615
Db        572 QSPPASHAHLQVAAA-PQSLCALRAVDQSVLLMKPEAELSVSSVYNLLTVKDLTNFPDNV 630

Qy        616 ---------------LYNTGYYLGMFMNS---FAVFQECGLWVLTDANLTKDYIDGVYDN 657
Db        631 DQQEEEQGHCPRPFFIHNGAIYVPLSSNEADIYSFLKGMGLKVFTNSKIRKPKSCSVIPS 690

Qy        658 AE-------------YAERFMEENEG--HIVDI-HDFSLGSSPH-VRKHFPETWIWLDT 699
Db        691 VSAGAVGQGYYGAGLGVVERPYVPQLGTYNVIPLNNEQSSGPVPETVRSYFPETWIWELV 750

Qy        700 NMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVI 759
Db        751 AVNSSGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISST-ASLRAFQPFFVELTMPYSVI 809

Qy        760 RGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVL 819
Db        810 RGEVFTLKATVLNYLPKCIRVSVQLKASPAF-----LASQNTKGEESYCICGSE-RQTLS 863

Qy        820 FPIRPTHLGEIPITVTALSPTA---------------SDAVTQMILVKAEGIEKSYSQS 863
Db        864 WTVTPKTLGNVNFSVSAEAMQSLELCGNEVVEVPEIKRKDTVIKTLLVEAEGIEQEKTFS 923

Qy        864 ILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGE 923
Db        924 ---SMTCASGANVSEQLSLKLPSNVVKESARASFSVLGDILGSAMQNIQNLLQMPYGCGE 980

Qy        924 QNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFG-NYDP 982
Db        981 QNMVLFAPNIYVLNYLNETQQLTQEIKAKAVGYLITGYQRQLNYKHQDGSYSTFGERYGR 1040

Qy        983 S-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG 1041
Db       1041 NQGNTWLTAFVLKTFAQARSYIFIDEAHITQSLTWLSQMQKDNGCFRSSGSLLNNAIKGG 1100

Qy       1042 NKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLES------EFSRGISDNYTLALITYA 1095
Db       1101 VEDEATLSAYVTIALLEIPLPVTNPIVRNALFCLESAWNVAKEGTHG-SHVYTKALLAYA 1159

Qy       1096 LSSVGSP-KAKEALNMLTWRAEQEGGMQFWVSSESKLSD-----SWQPRSLDIEVAAYAL 1149
Db       1160 FSLLGKQNQNREILNSLDKEAVKEDNLVHWERPQRPKAPVGHLYQTQAPSAEVEMTSYVL 1219

Qy       1150 LSHFLQFQTSEGIP----------IMRWLSRQRNSLGGFASTQDTTVALKALSEF--AAL 1197
Db       1220 LAYL----TAQPAPTSGDLTSATNIVKWIMKQQNAQGGFSSTQDTVVALHALSRYGAATF 1275

Qy       1198 MNTERTNIQVTVTGPSSPSP--------LAVVQPMAVN-------ISANGFGFAICQLNV 1242
Db       1276 TRTEKT-AQVTVQDSQTFSTNFQVDNNNLLLLQQISLPELPGEYVITVTGERCVYLQTSM 1334

Qy       1243 VYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLN--HVDLNVCTSFSGPGR---SGMA 1297
Db       1335 KYNILPE---------KEDSPFALKVQTVPQTCDGHKAHTSFQISLTISYTGNRPASNMV 1385

Qy       1298 LMEVNLLSGFMVPSEAISL---SETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKV 1354
Db       1386 IVDVKMVSGFIPLKPTVKMLERSSSVSRTEVSNNHVLIYVEQVTNQTLSFSFMVLQDIPV 1445

Qy       1355 SNTQDASVSIVDYYEPRRQAVRSY 1378
Db       1446 GDLKPAIVKVYDYYETDESVVAEY 1469
```

RESULT 12
A41185
alpha-2 macroglobulin MUG1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 12-Apr-1995
C;Accession: A41185
R;Overbergh, L.; Torrekens, S.; Van Leuven, F.; Van den Berghe, H.
J. Biol. Chem. 266, 16903-16910, 1991
A;Title: Molecular characterization of the murinoglobulins.
A;Reference number: A41185; MUID:91358495; PMID:1840592
A;Accession: A41185
A;Status: preliminary
A;Molecule type: mRNA

```
A;Residues: 1-1476 <OVE>
A;Cross-references: GB:M65736
C;Superfamily: alpha-2-macroglobulin

  Query Match             19.0%;  Score 1395.5;  DB 2;  Length 1476;
  Best Local Similarity   27.7%;  Pred. No. 7.4e-71;
  Matches  434;  Conservative  269;  Mismatches  555;  Indels  311;  Gaps   50;

Qy         13 LCVCTAALA-------VAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKT 65
              ||: :  ||         :     :::|  |              :|     | :: :     |
Db          9 LCLFSVLLAFLHSASLLNGDSKYMVLVPS------------QLYTETPEKICLHLYQLNE 56

Qy         66 ASNLTVSVLEAEG--------VFEKGSFKTLT--LPSLPLNSADEIYELRVTGRTQDEIL 115
                 :| |::    |          | :|   |: ::   :| |   :   ::     : : | |  :
Db         57 TVTVTASLVSQSGRKNLFDELVLDKDLFQCVSFIIPRLSSSDEEDFLYVDIKGPTHE--- 113

Qy        116 FSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNI-LIKDPKSN 174
              ||       :   :    :| |:||||  :|||  |  ||||:|::  :|      |   :  |:|||  |
Db        114 FSKRKAVLVKNKESVVFVQTDKPVYKPGQSVKFRVVSMDKMLRPLNELLPLAYIEDPKKN 173

Qy        175 LIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQV-QVNDQTYYQSFQVSEYVLPKFEVTL 233
               | ||     ::: :   |:    :|  |::  ||  |   |  |  :  :   :      || | |:|||:|  |
Db        174 RIMQWRDIKTENGLKQMSFSLAAEPIQGPYKIVVHKESGEKEEHSFTVMEFVLPRFNVDL 233

Qy        234 QTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFS 293
               : |    |:| :  |: |     ||||||||  |  |                          |||
Db        234 KVPNAMSVNDEVLSVTACGKYTYGKPVPGHV-------------------KINVCRETE 273

Qy        294 FNDEEMKNVMDSSNGLS----EYLDLSSPGP------VEILTTVTESVTGI--SRNVSTN 341
                  |: : :| :||  |           :| |          :   ||||   ||:   ||: :|
Db        274 TGCREVNSQLD-NNGCSTQEVNITELQSKKRNYEVQLFHVNATVTEEGTGLEFSRSGTTK 332

Qy        342 VFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEY 401
               :     :  |   |      :   :   :  |    |::      |:  :    |:       |  |      :  :||
Db        333 IERITNKLI--FLKADSHFRHGIPFFVKVRLVDIKGDPIPNEK----VFIKAQELSYTSA 386

Qy        402 WSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKA-----------YFLGS 450
                :      |                             |:    :   || |   :|                    |
Db        387 TTTDQHG----------------LAEFSIDTTCISGSS-LHIKVNHKEEDSCSYFYCMEE 429

Qy        451 KSSMAVHSLFK--SPSKTYIQLKTRDENI-------KVGSPFELVVSGNKR-LKEL--SY 498
               : :  |  |   :      | ||:||  |   |      :|                 |  :  |    ::  |:      ||||      |
Db        430 RHASAKHVAYAVYSLSKSYIYLDTETSSILPCNQIHTVQAHF--ILKGDLGVLKELIFYY 487

Qy        499 MVVSRGQLVAVGKQNSTM------------FSLTPENSWTPKACVIVYYIEDDGEIISDV 546
              :|::    ::     |        :                                  :   |  |    |   | :::|   |    |||:|:|
Db        488 LVMAHLSIIQTGNHTHQVEPGEAPVKGKFALEIPVEFSMVPMAKMLIYTILPDGEVIADS 547

Qy        547 LKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDIT 606
               :    ::     :||:  |  :|    ::  |:  :     |::: :  |:  |:  |||:||   |:          ::::
Db        548 VNFEIEKCLRNKVDLRFSTSQSLPASQTRLQVTAS-PQSLCGLRAVDQSVLLLKPESELS 606

Qy        607 MENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAE---- 662
                              :||      ||   |                    :      :     :       :    ::     ||
Db        607 PS------WIYNLP---GMQQNKF----------VPSSRLSEDQEDCILYSSWLAEKHTN 647

Qy        663 -----------RFMEEN-----------------EGHIVDIH----DFSLGSSPH----- 685
                            |::|:                                        :|   |       :|  |    :|
Db        648 LVPHGTEKDVYRYVEDMGLTAFTNLMIKLPIICFDYGMVPISAPRVEFDLAFTPEISWSL 707

Qy        686 -------------------------VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
                                       :||:||||||:|        :                 |  |:||||:||  |
Db        708 RTTLSKRPEEPPRKDPSSNDPLTETIRKYFPETWVWDIVTVNSTGLAEVEMTVPDTITEW 767

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
                  |        :|  |  ||||::   |  ||||:|||:  ::|||||:|||  |   |:  |:  |||    :  ::
Db        768 KAGALCLSNDTGLGLSSV-VPLQAFKPFFVEVSLPYSVVRGEAFMLKATVMNYLPTSMQM 826

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
                |  :|    |    |       :                |  :   :   :             |     :   :    |     ||   :     :|:|   :
Db        827 SVQLEASPDFTAVPVGDD------QDSYCLSANGRHTSSWLVTPKSLGNVNFSVSAEAQQ 880

Qy        841 AS---------------DAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSF 884
                  :|                  |  |  ::::|:  |||::  ::  |   |        :|    :      :  :|
Db        881 SSEPCGSEVATVPATGRKDTVVKVLIVEPEGIKQEHTFSSLFCASDAEIS---EKMSSGP 937

Qy        885 PPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQ 944
              ||   |    |    |           :  :||:|   :|           :   :  |||||||||:  ||||||:|   ||  :  :|
Db        938 PPTVVKDSARAHFSVMGDILSSAIRNTQNPLHMPYGCGEQNMVLFAPNIYVLKYLNETQQ 997

Qy        945 LTDNLKEKALSFMRQGYQRELLYQREDGSFSAFG--NYDPSGSTWLSAFVLRCFLEADPY 1002
              ||    :|  |||   |:|  ||||||    |  :||:||||     |  :   |:|||:||||:  |  :|  :
Db        998 LTQKIKTKALGFLRAGYQRELNYKHKDGSYSAFGDQNGEREGNTWLTAFVLKSFAQARAF 1057

Qy       1003 IDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKY 1062
               | ||:: :     :|||      ||  ||  |       | :  :: ::||        :||:|||   :||
Db       1058 IFIDESHITHAFTWLSQKQKDNGCFRSSGSLFNNAMKGGVDDEMTLSAYITMALLESSLP 1117

Qy       1063 QPNIDVQESIHFLES-----EFSRGISDNYTLALITYALSSVGSP-KAKEALNMLTWRAE 1116
                   :   | :::   |||              |     |      |     ||  ||:   ||   :      |:   |   |   |   |       |
Db       1118 ATHPVVSKALSCLESSWKTIEQERNASFVYTKALMAYAFALAGNQNKRDEILKSLDEEAI 1177

Qy       1117 QEGGMQFWV-SSESKLSDSW----QPRSLDIEVAAYALLSHFLQFQTSEGIP-------- 1163
               :|       |          :|:   |:                  |    |  ::|:  ||   :|:               |::   |
Db       1178 KENNSIHWKRPQKSRKSEHHLYKPQASSAEVEMNAYVVLARL----TAQPAPSPEDLTLS 1233

Qy       1164 --IMRWLSRQRNSLGGFASTQDTTVALKALSEFAAL---MNTERTNIQVTVTGPSSP--- 1215
                   :  ||::|:|| |||:|||||  |||  |||:  ||:     :    :   :   |  :  :   ||   |
Db       1234 MSTIMWLTKQQNSNGGFSSTQDTVVALDALSKYGAVTFSRSQKTTLVTIQSTGSFSQKFQ 1293

Qy       1216 ---SPLAVVQPMAV-------NISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFD 1265
                      |    ::|  :|:              ||  :|   |       |    :  ||:                               :     ||
Db       1294 VENSNRLLLQQVALPDIPGDYTISVSGEGCVYAQTMLRYNMHLE---------KQLSAFA 1344

Qy       1266 LDVAVK----ENKDDLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGFMVPSEAISLSETV 1320
               : |                  |       |      :::     |::|       |  |  :  :|  :|||||          |   |   ||
Db       1345 IWVQTVPLTCNNPKGHNSFQISLEISYTGSRPASNMVIADVKMLSGF------IPLKPTV 1398

Qy       1321 KKV---------EYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPR 1371
              ||:                |   :   :  :|||   |                :            ::     |    |    |     |     |    :    ||||
Db       1399 KKLERLEHVSRTEVSNNNVLIYLDQVTNQTLAFSFIIQQDIPVRNLQPAIVKVYDYYETD 1458

Qy       1372 RQAVRSYNS 1380
                 |    |:|
Db       1459 EMAFAEYSS 1467
```

RESULT 13
A20872
ovostatin precursor - chicken
N;Alternate names: ovomacroglobulin
C;Species: Gallus gallus (chicken)
C;Date: 05-Jun-1987 #sequence_revision 13-Mar-1997 #text_change 31-Mar-2000
C;Accession: I50671; I50672; S29836; A20872; A33715
R;Nielsen, K.L.; Sottrup-Jensen, L.; Nagase, H.; Thogersen, H.C.; Etzerodt, M.
DNA Seq. 5, 111-119, 1994
A;Title: Amino acid sequence of hen ovomacroglobulin (ovostatin) deduced from cloned cDN
A;Reference number: I50671; MUID:95218210; PMID:7535598
A;Accession: I50671
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1473 <NIE>
A;Cross-references: EMBL:X78801; NID:g671863; PIDN:CAA55384.1; PID:g671864
A;Accession: I50672
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 20-1473 <NI2>
A;Cross-references: EMBL:X78801; NID:g671863; PIDN:CAA55385.1; PID:g671865
R;Nielsen, K.L.; Sottrup-Jensen, L.
Biochim. Biophys. Acta 1162, 230-232, 1993
A;Title: Evidence from sequence analysis that hen egg-white ovomacroglobulin (ovostatin)
A;Reference number: S29836; MUID:93192299; PMID:7680577
A;Accession: S29836
A;Molecule type: protein
A;Residues: 977-980,'N',981-1007,'M',1009,'K',1010-1028 <NI3>
R;Nagase, H.; Harris Jr., E.D.; Woessner Jr., J.F.; Brew, K.

```
J. Biol. Chem. 258, 7481-7489, 1983
A;Title: Ovostatin: a novel proteinase inhibitor from chicken egg white. I. Purificatio
A;Reference number: A92427; MUID:83238315; PMID:6408074
A;Accession: A20872
A;Molecule type: protein
A;Residues: 37-49 <NAG>
R;Enghild, J.J.; Salvesen, G.; Brew, K.; Nagase, H.
J. Biol. Chem. 264, 8779-8785, 1989
A;Title: Interaction of human rheumatoid synovial collagenase (matrix metalloproteinase
 kinetics and identification of matrix metalloproteinase cleavage sites.
A;Reference number: A33715; MUID:89255343; PMID:2470748
A;Accession: A33715
A;Molecule type: protein
A;Residues: 710-717,'I',719,'T',721,'A',723-742,'L' <ENG>
C;Superfamily: alpha-2-macroglobulin
C;Keywords: alternative initiators; egg white; proteinase inhibitor

  Query Match             19.0%;  Score 1395;  DB 2;  Length 1473;
  Best Local Similarity   27.0%;  Pred. No. 7.8e-71;
  Matches  421;  Conservative  309;  Mismatches  522;  Indels  308;  Gaps   59;

Qy         18 AALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVS---VL 74
              ||     | |::::  | :::                ||||  ::    |    | |:|   ||
Db         33 AAAGKEPEPQYVLMVPAVLQ-----------SDSPSQVCLQFFNL----NQTISVRVVL 76

Qy         75 EAE----GVFEKGSFKTLTLPSL-----PLNSADEIY-ELRVTGRTQD-----EILFSNS 119
              | :      :|||  :  :   |   |     |: |    :      | | |       :: |
Db         77 EYDTINTTIFEKNTTTSNGLQCLNFMIPPVTSVSLAFISFTAKGTTFDLKERRSVMIWNM 136

Qy        120 TRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQ 178
                ||           ||:|||| :|||  |  |  ||:|  |   :|||  :       :: :|||:| |
Db        137 E--SF------VFVQTDKPIYKPGQSVMFRVVALDFNFKPVQEMYPLIAVQDPQNNRIFQ 188

Qy        179 WLSQQSDLGVISKTFQLSSHPILGDWSIQV-QVNDQTYYQSFQVSEYVLPKFEVTLQTPL 237
              | :   |:: :: |   |:  ||||:: |  |  : :  :    ||  |  ||||||||:||: |
Db        189 WQNVTSEINIVQIEFPLTEEPILGNYKIIVTKKSGERTSHSFLVEEYVLPKFDVTVTAPG 248

Qy        238 YCSMNSKHLNGTITAKYTYGKPVKGDVTLTFL-PLSFWGKKK------NITKTFKINGSA 290
                ::      |     |  |  ||||:|:|:|  |  |:             :|: |         :  ||     :|
Db        249 SLTVMDSELTVKICAVYTYGQPVEGKVQLSVCRDFDSYGRCKKSPVCQSFTKDLDTDGCL 308

Qy        291 NFSFNDE--EMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHD 348
              :   :  :  |:  :          |      ||          |: :  |   |  |   ::    |  ::   |
Db        309 SHILSSKVFELNRI-----GYKRNLD------VKAIVTEKEQVCNLTATQSISI--TQVM 355

Qy        349 YIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSG 408
                ::|  :     :   :   :|:    |  : :          :|  ||  :     |     :
Db        356 SSLQFENVDHHYRRGIPYFGQIKLVDKDNSPI------SNKVIQLFVNNKNTH------- 402

Qy        409 NQKMEAVQKINYTVPQSG--TFKIEFPILEDSSELQLKAYFLGS-------------KS 452
                              |:|   :|    |  |:    :    |   ||  |||  :   |
Db        403 ----------NFTTDINGIAPFSIDTSKIFD-PELSLKALYKTSDQCHSEGWIEPSYPDA 451

Qy        453 SMAVHSLFKSPSKTYIQLKTRDENIKVGSP------FELVVSGNKRLKELS--YMVVSRG 504
              |::|  |:  |  :  ::::::    :::   |               :  |    |  :  :  :    |:  :::|
Db        452 SLSVQRLY-SWTSSFVRIEPLWKDMSCGQKRMITVYYILNTEGYEHINIVNFYYVGMAKG 510

Qy        505 QLVAVGK--------QNST-MFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVF 555
              ::|  |:         || | |   |       |     ::|| :     |:::|·::  ::  |
Db        511 KIVLTGEIKVNIQADQNGTFMIPLVVNEKMAPALRLLVYMLHPAKELVADSVRFSIEKCF 570

Qy        556 KNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENV--VHE 613
              |||::|  :|:  :    :   |||  |       :|      :  |||||:  |:  :   :::  |  :  :|
Db        571 KNKVQLQFSEKQMLTTSNVSLVIEAA-ANSFCAVRAVDKSMLLLKSETELSAETIYNLHP 629

Qy        614 L------------------------ELYNTGYYL-----GMFMNSFAVFQECGLWVLTD 643
              :                        ::::  |  |        |:   :   :   ::  |:   |:
Db        630 IQDLQGYIFNGLNLEDDPQDPCVSSDDIFHKGLYYRPLTSGLGPDVYQFLRDMGMKFFTN 689

Qy        644 ANL------TKDYI---------------DGVYDNAEYAERFMEENEGHIVDIHDFSLGS 682
              : :        |:: :                    |  :||  |     |  :  ||::
Db        690 SKIRQPTVCTRETVRPPSYFLNAGFTASTHHVKLSAEVARE--ERGKRHILET------- 740

Qy        683 SPHVRKHFPETWIW---LDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTP 739
                :|:  ||||||   |   :  |               |:||:||  |  |:   :   |    |  |: :  |
Db        741 ---IREFFPETWIWDIILINSTG---KASVSYTIPDTITEWKASAFCVEELAGFGM-SVP 793

Qy        740 VELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEI 799
               | ||||||:  | ||||:| ||:|  :  :||||     ::  |::
Db        794 ATLTAFQPFFVDLTLPYSIIHGEDFLVRANVFNYLNHCIKINVLL-------------L 839

Qy        800 NATGHQQTLLVPSEDGA-------TVLFPIRPTHLGEIPITVTA--------------L 837
                :  :|   |: | :||       : ::  |  |   |::  ::||                :
Db        840 ESLDYQAKLISPEDDGCVCAKIRKSYVWNIFPKGTGDVLFSITAETNDDEACEEEALRNI 899

Qy        838 SPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQI 897
                |     : :||:  |||  :    :|:  |: :  |: :        : ::      |  |  |  ||  |
Db        900 RIDYRDTQIRALLVEPEGIRREETQNFLICMKDDVIS---QDVAIDLPTNVVEGSPRPSF 956

Qy        898 TAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFM 957
               : :||::|  :|   :    |::||:|  |||||:   ||||:||||  | :||::::|  |  :  ::
Db        957 SVVGDIMGTAIQNVHQLLQMPFGNGEQNMVLFAPNIYVLDYLDKTRQLSEDVKSKTIGYL 1016

Qy        958 RQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWL 1017
              |||::|  |:   :  |||:|  ||   |   |  |:|||:|||  :|:|  ::|  |  :|  :| | | |
Db       1017 VSGYQKQLSYKHPDGSYSTFGIRDKEGNTWLTAFVYKSFAEASRFIYIDDNVQAQTLIWL 1076

Qy       1018 KGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLES 1077
                  ||::|  |     |   ::::  ::|:|  ::  ::|:|||  :||          :  ::  :  :  ||:
Db       1077 ATKQKTDGCFQSTGILVNNAMKGGVENELSLSAYITIALLEAGHSMSHTVIRNAFYCLET 1136

Qy       1078 EFSRGISDNYTLALITYALSSVGSPKAKEA-LNMLTWRAEQEGGMQFWVSSE------SK 1130
                :  |:|  ||  ||:  ||    |   :  |:   |    |   |::   |::|  ::        |
Db       1137 ASEKNITDIYTQALVAYAFCLAGKAEICESFLRELQKSAKEVDGSKYWEQNQRSAPEKSH 1196

Qy       1131 LSDSWQPRSLDIEVAAYALLS-----HFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTT 1185
              |   |    |  |:|:  :|  ||:       :     |     |::|:  ||:||  ||||  |||
Db       1197 LLDHVQ--STDVEITSYVLLALLYKPNRSQEDLTKASAIVQWIIRQQNSYGGFASMQDTV 1254

Qy       1186 VALKALSEFAAL-----------MNTERTNIQV-TVTGPS----SPSPLAVVQPMAVNIS 1229
              |||:||:  :                           :|| :    :|   |   |  ||  |   :::
Db       1255 VALQALAAYGAATYNSVTQNVIKINSKNTFEKVFTVNNENRLLLQQTPLPQV-PGKYSLT 1313

Qy       1230 ANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKEN-----KDDLNHVDLNV 1284
                || |   :  |   : ||:                      :  ||    ::|:  :      :|       |:  :
Db       1314 VNGTGCVLIQTALRYNIHLP-----------EGAFGFSLSVQTSNASCPRDQPGKFDIVL 1362

Qy       1285 CTSFSGP-GRSGMALMEVNLLSGFMVPSEAISLSE-----TVKKVEYDHGKLNLYLDSV- 1337
                :|::|    |  |  :::|  :||||  ||  ::  ||  :        ||  :|||      :  |||  ::
Db       1363 ISSYTGKRSSSNMVIIDVKMLSGF-VPVKS-SLDQLIDDHTVMQVEYKKNHVLLYLGNIL 1420

Qy       1338 ----NETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDV 1393
                    |  |  |           ::|  |::  :   |   |   |  ||||      ||   |   |            ||   |
Db       1421 QKRRKEVTFSVE----QDFVVTHPKPAPVQIYDYYETEEYAVAEYMS--------LCRGV 1468

RESULT 14
A42210
alpha-1-macroglobulin precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C;Accession: A42210; A26124
R;Warmegard, B.; Martin, N.; Johansson, S.
Biochemistry 31, 2346-2352, 1992
A;Title: cDNA cloning and sequencing of rat alpha 1-macroglobulin.
A;Reference number: A42210; MUID:92172859; PMID:1371696
A;Accession: A42210
A;Molecule type: mRNA
A;Residues: 1-1500 <WAR>
A;Cross-references: GB:M84000; GB:J05359; NID:g205383; PIDN:AAA41591.1; PID:g205384
A;Experimental source: strain Sprague-Dawley, liver
A;Note: sequence extracted from NCBI backbone (NCBIP:87164)
R;Lonberg-Holm, K.; Reed, D.L.; Roberts, R.C.; Hebert, R.R.; Hillman, M.C.; Kutney, R.M.
J. Biol. Chem. 262, 438-445, 1987
A;Title: Three high molecular weight protease inhibitors of rat plasma. Isolation, charac
A;Reference number: A26124; MUID:87083487; PMID:2432067
```

```
A;Accession: A26124
A;Molecule type: protein
A;Residues: 25-29,'F',31-44 <LON>
C;Superfamily: alpha-2-macroglobulin
C;Keywords: glycoprotein; plasma; proteinase inhibitor
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-1474/Product: alpha-1-macroglobin #status predicted <MAT>

  Query Match             18.8%;  Score 1384;  DB 2;  Length 1500;
  Best Local Similarity   27.8%;  Pred. No. 3.4e-70;
  Matches  434;  Conservative  263;  Mismatches  588;  Indels  274;  Gaps   50;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVT--VKAELLKTASNL 69
Db         15 LLLLPRDATAATGKPRYVVLVPSELYAGVPEKVCVH-LNHLNETVTLNVTLEYGVQYSNL 73

Qy         70 TVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRI 129
Db         74 LID----QAVDKDSSYCSSFTISRPL-SPSALIAVEIKGPTHHFI----KKKSMWITKAE 124

Qy        130 S-VFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSDL- 186
Db        125 SPVFVQTDKPIYKPGQTVKFRVVSVDISFRPVNETFPVVYIENPKRNRIFQW--QNVDLP 182

Qy        187 -GVISKTFQLSSHPILGDWSIQVQVND-QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSK 244
Db        183 GGLHQLSFPLSVEPALGIYKVVVQKDSGKKIEHSFEVKEYVLPKFEVQVKMPKTMAFLEE 242

Qy        245 HLNGTITAKYTYGKPVKGDVTL------TFLPLSFWGK-KKNITKTFKINGSANFSFND- 296
Db        243 ELVVTACGLYTYGKPVPGLVTMKVCRKYTQSYSNCHGQHSKSICEEFSKQADEKGCFRQV 302

Qy        297 -----------------EEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVS 339
Db        303 VKTKVFQPRQKGYDMKIEVEAKIKEDGTG----IELTGTGSCEIANTLSK-----LKFTK 353

Qy        340 TNVFFKQHDYIIEFFDYTTVLK------PSLNFTATVKVTRADGNQLTLEERRNNVVITV 393
Db        354 ANTFYRPG---LPFFGQVLLVDEKGQPIPNKNLTVQVNSVRSQFTFTTDEHGLANILIDT 410

Qy        394 TQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSS 453
Db        411 T--NFTFSFMG--------------------------IRVIYKQNNICFDNWWVDEYHT 441

Qy        454 MAVHSLFK--SPSKTYIQLKTRDENIKVGSPFELVV------SGNKRLKELS--YMVVSR 503
Db        442 QADHSAARIFSPSRSYIQLELVLGTLACGQTQEIRIHFLLNEDALKDAKDLTFYYLIKAR 501

Qy        504 GQLVAVGK----------QNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQL 553
Db        502 GSIFNSGSHVFPLEQGKVKGVVSFPIRVEPGMAPVAKLIVYTILPNEELIADVQKFDIEK 561

Qy        554 VFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHE 613
Db        562 CFANTVNLSFPSAQSLPASDTHLTVKAT-PLSLCALTAVDQSVLLLKPEAKLSPQSIYNL 620

Qy        614 L-ELYNTGYYLGMF----------------------------MNSFAVFQECGLWVLTD 643
Db        621 LPQKAEQGAYLGPLPYKGGENCIKAEDITHNGIVYTPKQDLNDNDAYSVFQSIGLKIFTN 680

Qy        644 ANLTKDYIDGVYD--------------------NAEYAERFMEENEGHIVDIHDFSLGSS 683
Db        681 TRVHKPRYCPMYQAYPPLPYVGEPQALAMSAIPGAGY-------RSSNIRTSSMMMMGAS 733

Qy        684 ---------PHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLG 734
Db        734 EVAQEVEVRETVRKYFPETWIWDMVPLDLSGDGELPVKVPDTITEWKASAFCLSGTTGLG 793

Qy        735 LTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILM 794
Db        794 LSST-ISHKVFQPFFLELTLPYSVVRGEAFILKATVLNYMPHCIRIHVSLEMSPDFLAVP 852

Qy        795 TSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTA---LSP----------- 839
Db        853 VGS------HEDSHCICGNERKTVSWAVTPKSLGEVNFTATAEALQSPELCGNKVAEVPA 906

Qy        840 -TASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQIT 898
Db        907 LVQKDTVVKPVIVEPEGIEKEQTYNTLLCPQDAELQ---ENWTLDLPANVVEGSARATQS 963

Qy        899 AIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMR 958
Db        964 VLGDILGSAMQNLQNLLQMPYGCGEQNMVLFVPNIYVLEYLNETQQLTEAIKSKAISYLI 1023

Qy        959 QGYQRELLYQREDGSFSAFGN---YDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYT 1015
Db       1024 SGYQRQLNYQHSDGSYSTFGDRGMRHSQGNTWLTAFVLKAFAQAQSYIYIEKTHITNAFN 1083

Qy       1016 WLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFL 1075
Db       1084 WLSMKQRENGCFQQSGSLLNNAMKGGVDDEVTLSAYITIALLEMPLPVTHSVVRNALFCL 1143

Qy       1076 ESEFSRGISDN-----YTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFWVSSES 1129
Db       1144 ETAWA-SISNSQESHVYTKALLAYAFALAGNRAKRSEVLESLNKDAVNEEESVHWQRPKN 1202

Qy       1130 -------KLSDSWQPR--SLDIEVAAYALLSHFLQFQT-----------SEGIPIMRWLS 1169
Db       1203 VEENVREMRSFSYKPRAPSAEVEMTAYVLLAYLTSASSRPTRDLSSSDLTTASKIVKWIS 1262

Qy       1170 RQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQPMAVN-- 1227
Db       1263 KQQNSHGGFSSTQDTVVALQALSKYGAATFT-KSNKEVSVTIESSGT---VSGTLHVNNG 1318

Qy       1228 ------------------ISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDV- 1268
Db       1319 NRLLLQEVRLADLPGNYITKVSGSGCVYLQTSLKYNILPEA--------EGEAPFTLKVN 1370

Qy       1269 AVKENKDDLNH---VDLNVCTSFSGP-GRSGMALMEVNLLSGF--MVPS-EAISLSETVK 1321
Db       1371 TLPLNFDKAEHHRKFQIHINVSYIGERPNSNMVIVDVKMVSGFIPVKPSVKKLQDQSNIQ 1430

Qy       1322 KVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
Db       1431 RTEVNTNHVLIYIEKLTNQTMGFSFAVEQDIPVKNLKPAPVKVYDYYETDEFAIEEYSA 1489

RESULT 15
JC5144
murinoglobulin precursor - guinea pig
C;Species: Cavia porcellus (guinea pig)
C;Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 17-Mar-1999
C;Accession: JC5144
R;Iwasaki, H.; Suzuki, Y.; Sinohara, H.
J. Biochem. 120, 1167-1175, 1996
A;Title: Cloning and sequencing of cDNAs encoding plasma alpha-macroglobulin and murinog
A;Reference number: JC5143; MUID:97164019; PMID:9010766
A;Contents: liver
A;Accession: JC5144
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1464 <IWA>
A;Cross-references: DDBJ:D84339
C;Superfamily: alpha-2-macroglobulin
F;1-23/Domain: signal sequence #status predicted <SIG>

  Query Match             18.7%;  Score 1374.5;  DB 2;  Length 1464;
  Best Local Similarity   27.9%;  Pred. No. 1.1e-69;
  Matches  430;  Conservative  243;  Mismatches  596;  Indels  271;  Gaps   49;

Qy         13 LCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVS 72
Db         16 LALLMTAASVHKKPQYMVLVP------------TQLYTEVPEKSCLHLHYLKETVTVSAS 63

Qy         73 VLEA-------------EGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNS 119
```

```
Db         64 LMSSMGKKSLFSDFEVHEDLFQCVSF---TLPRTSSPNEVAFLSVQIKGRTH---TFSEE 117

Qy        120 TRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKP-YKTSLNILIKDPKSNLIQQ 178

Db        118 MAVVVRNTESILIVQTDKPMYKPGQTVKFRVVSVDRNLRQRHELFPLVYIEDPRKNRIMQ 177

Qy        179 WLSQQSDLGVISKTFQLSSHPILGDWSIQV-QVNDQTYYQSFQVSEYVLPKFEVTLQTPL 237

Db        178 WRDITSENGIKQLSFNLSSEPIQGFYKIMVLKKSGEKKEHVFSAEEYVLPRFEVQVKCPK 237

Qy        238 YCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFN-- 295

Db        238 AVTFLDEKVNVTVCGKYTYGKPVLGHV------------KLNICHLFKSYFGSEYGHNYG 285

Qy        296 ----DEEMKNVMDSSNGLSEYL--------DLSSPGPVEILTTVTESVTGIS------- 335

Db        286 GQDPCEEISQQLNSQGCITQAVRTSRDFMNWNMSHIFNLQVNAKITEEGTGLEFTGTGTT 345

Qy        336 --RNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITV 393

Db        346 EVRKPRTNLVRVSTD---------SHFRPGIPFFVVQFRLEDVKG----VPVADKHIFLKV 392

Qy        394 TQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKI----------EFPILEDSSELQL 443

Db        393 FTTKYNS--SATTDEHGLVEFSTNTTDIVAPSLTFKVYYKEQGHCFKHYCIEEELKEVDF 450

Qy        444 KAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVS---GNKRLKELS--- 497

Db        451 VFY----------SVF-SYSKSFVYLEPVTGALPCGQMHTVHVHYTLNGKVLGELREMV 498

Qy        498 --YMVVSRGQLVAVGKQNSTM----------FSLTPENSWTPKACVIVYYIEDDGEIISD 545

Db        499 FYYLIMARGIIVQTGTHAFFMEPGELKGYFNLSIPIESYMAPAPQMLIYAILPSGEVIAD 558

Qy        546 VLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDI 605

Db        559 SAKFEIENCLLSQVGLSFRPAQSLPASQTHLRVT-APPQSLCALRAVDQSVLLMRPEAEL 617

Qy        606 TMENVVHELEL----------------------YNTGYYLGMFMN---SFAVFQECGLW 639

Db        618 SPSSIYNLLSIKDTDSSSDLNWLANFPENCKDINYDKIKYSSLHRNDKDTFKFIESMGLN 677

Qy        640 VLT-----DANLTKDYIDGVYDNA---EYAERFMEENEGHIVDIHDFSLGSSPHVRKHFP 691

Db        678 IFTNLKIRDPNLCPD--DLLFPQAVAIKYTD--------HTVPAEVDEKPLSETVRTYFP 727

Qy        692 ETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIF 751

Db        728 ETWIWDLVEVNSSGVAEVAVTVPDTITEWKAGALCLPSDTGLGLSPT-ASLRAFQPFFVE 786

Qy        752 LNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVP 811

Db        787 LTMPYSVVRGEVFTFKATVLNYLSKCIQVSVQLEASPAFTAVPLAKD-------QDSYCL 839

Qy        812 SEDG-ATVLFPIRPTHLGEIPITVTALSPTAS----------------DAVTQMILVKAE 854

Db        840 CEDGRQTVSWLVTPKALGNVNFSVSAETQQPSELCGTEVAAVPEARKKDTVVKPLLVEPE 899

Qy        855 GIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL 914

Db        900 GIKKEYMINSFLCASDAVIS---EKLSLKLPPKLVTDSARAFFSVFGDILSSSIKNTQNL 956

Qy        915 IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSF 974

Db        957 LQMPYGCGEQNMVLFAPNIYVLKYLNETDQLTQEIKSKAIGYLSAGYQRQLNYKHPDGSY 1016

Qy        975 SAFG--NYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGR 1032

Db       1017 SAFGHQNGRNQGNTWLTAFVLKTLAQARAFIFIDETHISHAFTWLSQKQKDNGCFRSSGT 1076

Qy       1033 VIHSELQGGNKSPVTLTAYIVTSLL--GYRKYQPNID-----VQESIHFLESEFSRGISD 1085

Db       1077 LFHNDLKGGVDDEVTLSAYITIALLEMPLPATHPVVSKALTCLNESLE-MAKEGSNG-SY 1134

Qy       1086 NYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFWVSSESKLSDS---WQPR--S 1139

Db       1135 VYTKALLAYAFALAGNQDKRKEILKSLDEEAVKEDNSIHWERPQKPRRPEALLYQPQAPS 1194

Qy       1140 LDIEVAAYALLSHFLQFQTSEGIP----------IMRWLSRQRNSLGGFASTQDTTVALK 1189

Db       1195 AEVEMTSYVLLAHL----TAQGTPTPEEMTSAMRIVNWITKQQNSYGGFSSTQDTVVALH 1250

Qy       1190 ALSEFAAL---------------MNTERTNIQVTVTGPSSPSPLAVVQ------PMAVNI 1228

Db       1251 ALSRYGAATFGRTGKTALVKIQSSGTFSTKFQV-----DNSNRLLLQQVSLPHIPEEYTI 1305

Qy       1229 SANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDV-AVKENKDDLN-HVDLNVCT 1286

Db       1306 SVSGEGCVYTQTALKYNVFLE---------KEKYAFALQVHTVPQTCDDPKAHKRFQISL 1356

Qy       1287 SFSGPGR---SGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHGKLNLYLDSVNET 1340

Db       1357 EVSYTGRRPASNMAIVNVKMISGFSPLKSTVKMLERSDHVSRTEVNKNNVLIYVDQVTNE 1416

Qy       1341 QFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380

Db       1417 TLNFSFDVLQDVPVRNLKPAFVKVYDYYETGEAAFAEYSA 1456
```

Search completed: January 15, 2004, 18:14:44
Job time : 39 secs

```
                       GenCore version 5.1.6
              Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:        January 15, 2004, 18:03:24 ; Search time 18 Seconds
                                            (without alignments)
                                            3730.785 Million cell updates/sec

Title:         US-10-020-095-4
Perfect score: 7348
Sequence:      1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table: BLOSUM62
               Gapop 10.0 , Gapext 0.5

Searched:      127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters:      127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     SwissProt_41:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result            Query
  No.    Score    Match Length DB ID                        Description
-----------------------------------------------------------------------------
    1     1455    19.8   1477  1  A1I3_RAT                  P14046 rattus norv
    2     1440    19.6   1474  1  A2MG_HUMAN                P01023 homo sapien
    3   1412.5    19.2   1472  1  A2MG_RAT                  P06238 rattus norv
    4   1406.5    19.1   1476  1  A2M1_MOUSE                P28665 mus musculu
    5     1402    19.1   1451  1  A2M2_MOUSE                P28666 mus musculu
    6   1396.5    19.0   1482  1  PZP_HUMAN                 P20742 homo sapien
    7     1395    19.0   1473  1  OVOS_CHICK                P20740 gallus gall
    8   1373.5    18.7   1495  1  A2MG_MOUSE                Q61838 mus musculu
    9   1026.5    14.0   1673  1  CO3_LAMJA                 Q00685 lampetra ja
   10    991.5    13.5   1744  1  CO4_HUMAN                 P01028 homo sapien
   11      987    13.4   1738  1  CO4_MOUSE                 P01029 mus musculu
   12    984.5    13.4   1640  1  CO3_ONCMY                 P98093 oncorhynchu
   13    973.5    13.2   1651  1  CO3_NAJNA                 Q01833 naja naja (
   14      966    13.1   1663  1  CO3_RAT                   P01026 rattus norv
   15      956    13.0   1663  1  CO3_MOUSE                 P01027 mus musculu
   16      950    12.9   1737  1  CO4_RAT                   P08649 rattus norv
   17      934    12.7   1663  1  CO3_HUMAN                 P01024 homo sapien
   18      921    12.5   1676  1  CO5_HUMAN                 P01031 homo sapien
   19      905    12.3   1680  1  CO5_MOUSE                 P06684 mus musculu
   20      904    12.3   1666  1  CO3_CAVPO                 P12387 cavia porce
   21    857.5    11.7   1620  1  CO3_EPTBU                 P98094 eptatretus
   22      543     7.4    920  1  CO4_BOVIN                 P01030 bos taurus
   23    467.5     6.4   1536  1  Y984_THEMA                Q9x079 thermotoga
   24    432.5     5.9    726  1  CO3_RABIT                 P12247 oryctolagus
   25    290.5     4.0   1892  1  Y835_RICCN                Q92hd6 rickettsia
   26      283     3.9   1653  1  YFHM_ECOLI                P76578 escherichia
   27    263.5     3.6   1906  1  YFA0_ANASP                Q8ym40 anabaena sp
   28      227     3.1   1582  1  YU30_RALSO                Q8xv02 ralstonia s
   29      187     2.5   1516  1  Y8I9_PSEAE                Q9hvt2 pseudomonas
   30      183     2.5   1534  1  YFAS_ECO57                Q8xe35 escherichia
   31    180.5     2.5   1720  1  FTSH_CHLVU                P56369 chlorella v
   32      176     2.4   1534  1  YFAS_ECOLI                P76464 escherichia
   33    174.5     2.4   2366  1  TOXB_CLODI                P18177 clostridium
   34    173.5     2.4   1905  1  Y659_PASMU                Q9cmz1 pasteurella
   35    166.5     2.3   2358  1  YEEJ_ECOLI                P76347 escherichia
   36    164.5     2.2   2004  1  YP73_YERPE                Q8zdj2 yersinia pe
   37      161     2.2   1337  1  P152_YEAST                P39685 saccharomyc
   38      161     2.2   1562  1  YM81_YEAST                Q04781 saccharomyc
   39      161     2.2   2660  1  YEEJ_ECO57                Q8x8v7 escherichia
   40      161     2.2   3354  1  CADN_MOUSE                Q99pf4 mus musculu
   41      161     2.2   4349  1  FAT2_HUMAN                Q9nyq8 homo sapien
   42    159.5     2.2   1609  1  FIG2_YEAST                P25653 saccharomyc
   43      159     2.2   1656  1  ATC8_YEAST                Q12674 saccharomyc
   44      159     2.2   2201  1  TENA_HUMAN                P24821 homo sapien
   45      159     2.2   3174  1  CHAC_HUMAN                Q96r17 homo sapien

                              ALIGNMENTS

RESULT 1
A1I3_RAT
ID   A1I3_RAT      STANDARD;      PRT;  1477 AA.
AC   P14046;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha-1-inhibitor III precursor.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=88153707; PubMed=2831216;
RA   Braciak T.A., Northemann W., Hudson G.O., Shiels B.R., Gehring M.R.,
RA   Fey G.H.;
RT   "Sequence and acute phase regulation of rat alpha 1-inhibitor III
RT   messenger RNA.";
RL   J. Biol. Chem. 263:3999-4012(1988).
CC   -!- FUNCTION: PROTEASE INHIBITOR WITH A WIDE SPECTRUM OF PROTEIN
CC       TARGETS, WHICH ATTACHES THROUGH ITS THIOLESTER FUNCTION.
CC   -!- SUBUNIT: Monomer.
CC   -!- MISCELLANEOUS: THIS INHIBITOR HAS REDUCED PLASMA CONCENTRATIONS IN
CC       ACUTE PHASE.
CC   -!- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC       INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J03552; AAA40628.1; -.
DR   EMBL; M22358; AAA40629.1; -.
DR   PIR; A29952; A29952.
DR   HSSP; P01023; 1BV8.
DR   GO; GO:0017114; F:wide-spectrum protease inhibitor activity; NAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW   Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
KW   Liver; Thioester bond.
FT   SIGNAL          1     24
FT   CHAIN          25   1477       ALPHA-1-INHIBITOR III.
FT   DOMAIN        601    750       BAIT REGION (APPROXIMATELY).
FT   DISULFID       48     86       BY SIMILARITY.
FT   DISULFID      251    295       BY SIMILARITY.
FT   DISULFID      269    283       BY SIMILARITY.
FT   DISULFID      468    563       BY SIMILARITY.
```

```
FT   DISULFID    595    774       BY SIMILARITY.
FT   DISULFID    643    678       BY SIMILARITY.
FT   DISULFID    850    886       BY SIMILARITY.
FT   DISULFID    924   1324       BY SIMILARITY.
FT   DISULFID   1082   1130       BY SIMILARITY.
FT   DISULFID   1355   1470       BY SIMILARITY.
FT   CROSSLNK    975    978       Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD     55     55       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    247    247       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    301    301       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    321    321       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    393    393       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    508    508       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    750    750       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    777    777       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    872    872       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    994    994       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1143   1143       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1314   1314       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1427   1427       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT       4      4       D -> N.
FT   VARIANT      87     87       V -> L.
FT   VARIANT     137    137       M -> V.
SQ   SEQUENCE   1477 AA;  163773 MW;  4DC05367C8385D2B CRC64;

  Query Match             19.8%;  Score 1455;  DB 1;  Length 1477;
  Best Local Similarity   29.4%;  Pred. No. 8.4e-72;
  Matches  457;  Conservative  237;  Mismatches  578;  Indels  284;  Gaps   50;

Qy         13 LCVCTAALAVAPGPRFLVTAPGIIRPGGN----VTIGVELLEHCPSQVTVKAELLKTASN 68
Db          9 LCLFSALLAFLPFASLL---------NGNSKYMVLVPSQLYTETPEKICLHLYHLNETVT 59

Qy         69 LTVSVLEAEG--------VFEKGSFK--TLTLPSLPLNSADEIYELRVTGRTQDEILFSN 118
Db         60 VTASLISQRGTRKLFDELVVDKDLFHCVSFTIPRLPSSEEEESLDINIEGAKHK---FSE 116

Qy        119 STRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-----IKDPKS 173
Db        117 RRVVLVKNKESVVFVQTDKPMYKPGQSVKFRVVSMDKNLHP----LNELFPLAYIEDPKM 172

Qy        174 NLIQQWLSQQSDLGVISKTFQLSSHPILGDWSI----QVQVNDQTYYQSFQVSEYVLPKF 229
Db        173 NRIMQWQDVKTENGLKQLSFSLSAEPIQGPYKIVILKQSGVKEE---HSFTVMEFVLPRF 229

Qy        230 EVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGS 289
Db        230 GVDVKVPNAISVYDEIINVTACATYTYGKPVPGHVKISL-----------------CHGN 272

Qy        290 ANFSFNDEEMKNVMDS---SNGLS------------EYLDLSSPGPVEILTTVTESVTG 333
Db        273 PTFSSETKSGCKEEDSRLDNNGCSTQEVNITEFQLKENYLKMHQ--AFHVNATVTEEGTG 330

Qy        334 ISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITV 393
Db        331 SEFSGSGRIEVERTRNKFLFLKADSHFRHGIPFFVKVRLVDIKGDPIPNEQ----VLIKA 386

Qy        394 TQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSS 453
Db        387 RDAGYTNATTTDQHGLAK----------------FSIDTNGISDYS-LNIKVYH--KEES 427

Qy        454 MAVHS---------------LFKSPSKTYIQLKTR------DENIKVGSPFEL--VVSGN 490
Db        428 SCIHSSCTAERHAEAHHTAYAVYSLSKSYIYLDTEAGVLPCNQIHTVQAHFILKGQVLGV 487

Qy        491 KRLKELSYMVVSRGQLVAVGK-----------QNSTMFSLTPENSWTPKACVIVYYIED 538
Db        488 LQQIVFHYLVMAQGSILQTGNHTHQVEPGESQVQGNFALEIPVEFSMVPVAKMLIYTILP 547

Qy        539 DGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNL 598
Db        548 DGEVIADSVKFQVEKCLRNKVHLSFSPSQSLPASQTHMRVTAS-PQSLCGLRAVDQSVLL 606

Qy        599 ----------------MNASNDITMENVVHELELYNTGYY-LGMFMNSFAVFQECGLWV 640
Db        607 QKPEAELSPSLIYDLPGMQDSNFIASSNDPFEDEDYCLMYQPIAREKDVYRYVRETGLMA 666

Qy        641 LTDANL-TKDYIDGVYDNAEYA----------ERFMEENEGHIVDIHDFSLGSSP----- 684
Db        667 FTNLKIKLPTYCNTDYDMVPLAVPAVALDSSTDRGMYESL-PVVAVKS-PLPQEPPRKDP 724

Qy        685 --------HVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLT 736
Db        725 PPKDPVIETIRNYFPETWIWDLVTVNSSGVTELEMTVPDTITEWKAGALCLSNDTGLGLS 784

Qy        737 TTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTS 796
Db        785 SV-ASFQAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPTSLPMAVLLEASPDF------ 837

Qy        797 SEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----PTAS--------- 842
Db        838 TAVPVENNQDSYCLGANGRHTSSWLVTPKSLGNVNFSVSAEARQSPGPCGSEVATVPETG 897

Qy        843 --DAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
Db        898 RKDTVVKVLIVEPEGIKKEHTFSSLLCASDAELS---ETLSLLLPPTVVKDSARAHFSVM 954

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
Db        955 GDILSSAIKNTQNLIQMPYGCGEQNMVLFAPNIYVLKYLNETQQLTEKIKSKALGYLRAG 1014

Qy        961 YQRELLYQREDGSFSAFGNYD--PSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLK 1018
Db       1015 YQRELNYKHKDGSYSAFGDHNGQGQGNTWLTAFVLKSFAQARAFIFIDESHITDAFTWLS 1074

Qy       1019 GHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNID--VQESIHFLE 1076
Db       1075 KQQKDSGCFRSSGSLLNNAMKGGVDDEITLSAYITMALL--ESSLPDTDPVVSKALSCLE 1132

Qy       1077 SEFSR-----GISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW------ 1124
Db       1133 SSWENIEQGGNGSFVYTKALMAYAFALAGNQEKRNEILKSLDKEAIKEDNSIHWERPQKP 1192

Qy       1125 VSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIP----------IMRWLSRQRNS 1174
Db       1193 TKSEGYLYTP-QASSAEVEMSAYVVLARL----TAQPAPSPEDLALSMGTIKWLTKQQNS 1247

Qy       1175 LGGFASTQDTTVALKALSEF-AALMNTERTNIQVTVTGPSSPSPLAVVQ----------- 1222
Db       1248 YGGFSSTQDTVVALDALSKYGAATFSKSQKTPSVTVQSSGSFSQKFQVDKSNRLLLQQVS 1307

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK----ENK 1274
Db       1308 LPYIPGNYTVSVSGEGCVYAQTTLRYNVPLE---------KQQPAFALKVQTVPLTCNNP 1358

Qy       1275 DDLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGFMVPSEAISLSETVKKV----------E 1324
Db       1359 KGQNSFQISLEISYMGSRPASNMVIADVKMLSGF------IPLKPTVKKLERLGHVSRTE 1412

Qy       1325 YDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
Db       1413 VTTNNVLLYLDQVTNQTLSFSFIIQQDIPVKNLQPAIVKVYDYYETDEVAFAEYSS 1468

RESULT 2
A2MG_HUMAN
ID   A2MG_HUMAN     STANDARD;      PRT;  1474 AA.
AC   P01023; Q13677;
DT   21-JUL-1986 (Rel. 01, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha-2-macroglobulin precursor (Alpha-2-M).
GN   A2M.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85190481; PubMed=2581245;
RA   Kan C.-C., Solomon E., Belt K.T., Chain A.C., Hiorns L.R., Fey G.H.;
RT   "Nucleotide sequence of cDNA encoding human alpha 2-macroglobulin and
RT   assignment of the chromosomal locus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:2282-2286(1985).
RN   [2]
RP   SEQUENCE FROM N.A., AND VARIANT HIS-704.
RX   MEDLINE=92246939; PubMed=1374237;
RA   Matthijs G., Devriendt K., Cassiman J.-J., van den Berghe H.,
RA   Marynen P.;
RT   "Structure of the human alpha-2 macroglobulin gene and its promotor.";
RL   Biochem. Biophys. Res. Commun. 184:596-603(1992).
RN   [3]
RP   SEQUENCE OF 24-1474, AND CHARACTERIZATION.
RX   MEDLINE=84239807; PubMed=6203908;
RA   Sottrup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA   Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RT   "Primary structure of human alpha 2-macroglobulin. V. The complete
RT   structure.";
RL   J. Biol. Chem. 259:8318-8327(1984).
RN   [4]
RP   ERRATUM.
RA   Sottrup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,
RA   Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;
RL   J. Biol. Chem. 260:6500-6500(1985).
RN   [5]
RP   SEQUENCE OF 832-1474 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=85219061; PubMed=2408344;
RA   Bell G.I., Rall L.B., Sanchez-Pescador R., Merryweather J.P.,
RA   Scott J., Eddy R.L., Shows T.B.;
RT   "Human alpha 2-macroglobulin gene is located on chromosome 12.";
RL   Somat. Cell Mol. Genet. 11:285-289(1985).
RN   [6]
RP   SEQUENCE OF 672-747.
RX   MEDLINE=90242963; PubMed=1692292;
RA   Marynen P., Devriendt K., van den Berghe H., Cassiman J.J.;
RT   "A genetic polymorphism in a functional domain of human pregnancy
RT   zone protein: the bait region. Genomic structure of the bait domains
RT   of human pregnancy zone protein and alpha 2 macroglobulin.";
RL   FEBS Lett. 262:349-352(1990).
RN   [7]
RP   INHIBITORY SITE.
RX   MEDLINE=84030513; PubMed=6195065;
RA   Virca G.D., Salvesen G.S., Travis J.;
RT   "Human neutrophil elastase and cathepsin G cleavage sites in the bait
RT   region of alpha 2-macroglobulin. Proposed structural limits of the
RT   bait region.";
RL   Hoppe-Seyler's Z. Physiol. Chem. 364:1297-1302(1983).
RN   [8]
RP   INHIBITORY SITE.
RX   MEDLINE=81212827; PubMed=6165619;
RA   Sottrup-Jensen L., Loenblad P.B., Stepanik T.M., Petersen T.E.,
RA   Magnusson S., Joernvall H.;
RT   "Primary structure of the 'bait' region for proteinases in alpha 2-
RT   macroglobulin. Nature of the complex.";
RL   FEBS Lett. 127:167-173(1981).
RN   [9]
RP   INHIBITORY SITE.
RX   MEDLINE=81255805; PubMed=6167263;
RA   Hall P.K., Nelles L.P., Travis J., Roberts R.C.;
RT   "Proteolytic cleavage sites on alpha 2-macroglobulin resulting in
RT   proteinase binding are different for trypsin and Staphylococcus
RT   aureus V-8 proteinase.";
RL   Biochem. Biophys. Res. Commun. 100:8-16(1981).
RN   [10]
RP   INHIBITORY SITE.
RX   MEDLINE=82095610; PubMed=6172288;
RA   Mortensen S.B., Sottrup-Jensen L., Hansen H.F., Petersen T.E.,
RA   Magnusson S.;
RT   "Primary and secondary cleavage sites in the bait region of alpha 2-
RT   macroglobulin.";
RL   FEBS Lett. 135:295-300(1981).
RN   [11]
RP   STRUCTURE BY NMR OF 1337-1474.
RX   MEDLINE=99081553; PubMed=9865955;
RA   Huang W., Dolmer K., Liao X., Gettins P.G.W.;
RT   "Localization of basic residues required for receptor binding to the
RT   single alpha-helix of the receptor binding domain of human alpha2-
RT   macroglobulin.";
RL   Protein Sci. 7:2602-2612(1998).
RN   [12]
RP   VARIANT ILE-1000.
RX   MEDLINE=91187639; PubMed=1707161;
RA   Poller W., Faber J.-P., Olek K.;
RT   "Sequence polymorphism in the human alpha2-macroglobulin (A2M) gene.";
RL   Nucleic Acids Res. 19:198-198(1991).
RN   [13]
RP   VARIANTS TYR-972 AND ILE-1000.
RX   MEDLINE=92128897; PubMed=1370808;
RA   Poller W., Faber J.-P., Klobeck G., Olek K.;
RT   "Cloning of the human alpha 2-macroglobulin gene and detection of
RT   mutations in two functional domains: the bait region and the
RT   thiolester site.";
RL   Hum. Genet. 88:313-319(1992).
CC   -!- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A
CC       UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,
CC       CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES
CC       FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT
CC       REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH
CC       TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST
CC       LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR
CC       WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE
CC       BAIT REGION A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE
CC       COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.
CC   -!- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-
CC       LINKED CHAINS.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- MISCELLANEOUS: CONTRARY TO THE RAT PROTEIN, WHICH IS AN ACUTE
CC       PHASE PROTEIN, THIS PROTEIN IS ALWAYS PRESENT AT HIGH LEVELS IN
CC       CIRCULATION.
CC   -!- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC       INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M11313; AAA51551.1; -.
DR   EMBL; Z11711; CAA77774.1; -.
DR   EMBL; M36501; AAA51552.1; -.
DR   EMBL; X68728; CAA48670.1; -.
DR   EMBL; X68729; CAA48670.1; JOINED.
DR   PIR; A94033; MAHU.
DR   PDB; 1BV8; 30-SEP-98.
DR   SWISS-2DPAGE; P01023; HUMAN.
DR   Genew; HGNC:7; A2M.
DR   MIM; 103950; -.
DR   GO; GO:0016975; F:alpha-2 macroglobulin; NAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; MacroglobInA2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW   Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
KW   3D-structure; Polymorphism; Thioester bond.
FT   SIGNAL        1     23
FT   CHAIN        24   1474       ALPHA-2-MACROGLOBULIN.
FT   SITE        693    694       TRANSGLUTAMINASE CROSS-LINKING SITE
```

```
FT                                     (POTENTIAL).
FT   DOMAIN       690    728           BAIT REGION.
FT   SITE         704    709           INHIBITORY.
FT   SITE         719    723           INHIBITORY.
FT   SITE         730    735           INHIBITORY.
FT   DISULFID      48     86
FT   DISULFID     251    299
FT   DISULFID     269    287
FT   DISULFID     278    431
FT   DISULFID     470    470           INTERCHAIN (WITH OTHER CHAIN).
FT   DISULFID     563    563           INTERCHAIN (WITH OTHER CHAIN).
FT   DISULFID     595    771
FT   DISULFID     642    689
FT   DISULFID     821    849
FT   DISULFID     847    883
FT   DISULFID     921   1321
FT   DISULFID    1079   1127
FT   DISULFID    1352   1467
FT   CROSSLNK     972    975           Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD      55     55           N-LINKED (GLCNAC. . .).
FT   CARBOHYD      70     70           N-LINKED (GLCNAC. . .).
FT   CARBOHYD     247    247           N-LINKED (GLCNAC. . .).
FT   CARBOHYD     396    396           N-LINKED (GLCNAC. . .).
FT   CARBOHYD     410    410           N-LINKED (GLCNAC. . .).
FT   CARBOHYD     869    869           N-LINKED (GLCNAC. . .).
FT   CARBOHYD     991    991           N-LINKED (GLCNAC. . .).
FT   CARBOHYD    1424   1424           N-LINKED (GLCNAC. . .).
FT   VARIANT      704    704           R -> H (IN dbSNP:1800434).
FT                                     /FTId=VAR_000012.
FT   VARIANT      972    972           C -> Y (PROBABLY INTERFERES WITH THE
FT                                     ACTIVITY; dbSNP:1800433).
FT                                     /FTId=VAR_000013.
FT   VARIANT     1000   1000           V -> I (IN dbSNP:669).
FT                                     /FTId=VAR_000014.
FT   CONFLICT      63     63           MISSING (IN REF. 3).
FT   CONFLICT     563    563           C -> E (IN REF. 3).
FT   CONFLICT    1148   1148           A -> D (IN REF. 5).
FT   CONFLICT    1195   1195           H -> D (IN REF. 5).
FT   STRAND      1341   1347
FT   TURN        1352   1353
FT   HELIX       1355   1359
FT   STRAND      1360   1369
FT   STRAND      1379   1384
FT   STRAND      1390   1391

  Query Match             19.6%;  Score 1440;  DB 1;  Length 1474;
  Best Local Similarity   29.1%;  Pred. No. 5.6e-71;
  Matches  451;  Conservative  266;  Mismatches  551;  Indels  284;  Gaps   59;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
              || :      :|:  |::|  | ::           | ||  :     ||| |       |::
Db         15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy         72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
               :  ||||   |      :|          ::|  :|::|       ::| |  ||:    |     | :  :
Db         74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy        127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
               :    ||:|:|||::|||  | ||||:|::   :| |      : :: |:|||  |  |  ||  |  |  :
Db        125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy        186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
               |:     :|  |||  |    | :  :  ||          :|  :   |  |  |:||||||  :  |      ::
Db        185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy        244 KHLNGTITAKYTYGKPVKGDVTLTFLPL--------------------------SFWGK 276
               :  :|  ::         ||||||  |  ||::                               |:  :
Db        244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303

Qy        277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
               |    ||  |::            :  |||          |      |    |    ||          :|  :
Db        304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy        332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
               | :|   |    :   |:|    | ||                       :|    ||    : :         | ||
Db        348 TKLS-FVKVDSHFRQG---IPFFG---------------QVRLVDGKGVPIP----NKVI 384

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
               :  :   |   |:  ||:   :     ||  ||  |   :  |  |  :: :                :  |:
Db        385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441

Qy        439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
               |   .  ||               |   |||:::  |:         :    |               :::|
Db        442 EEAHHTAY-----------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489

Qy        493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
               ||:||  |::::|  :|  |            ||            |:  ::    |  |  :::|  :   |
Db        490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
               ::|  |    |:      ||:  |  :|    :: |:         ||::         |  |:    :  |||:||  ||
Db        550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608

Qy        601 ASNDITMENVVHEL-ELYNTGY--------------YLGMFMNS-------------FA 631
               :::    :|  :  |  |     ||:                   :::|
Db        609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 668

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAE----------YAERFMEENEGHIVDIHDFSL 680
               :: ||    |:: :  |    :              |              |           |           :|  :  :
Db        669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725

Qy        681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
                   ||    |||:||||||        :           |     |||||:||  |  |     |  :|||  |||:::
Db        726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782

Qy        738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
               |    |:|||||:  |  :|||||||  |  |:  |:  |||         |  |  :|  |    |  :
Db        783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841

Qy        798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
               |         |  |  |  |                |  |  :  :  |   ||:|  :
Db        842 EQAPHCICANGRQ-----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890

Qy        842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
                       |  |  :  :||:  ||:||     :  :  ||    :       :     :  ||       |||  |     |  |
Db        891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947

Qy        897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
               :: :||:||  ::        :|::||||||||||: ||||||:||||  :  :|||    :|  ||:  :
Db        948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEVKSKAIGY 1007

Qy        957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
               :   ||||:|  |:       |||:|  ||          |        :  |:|||:||||:  |  :|       ||  ||:    :   :
Db       1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067

Qy       1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
               ||    ||    ||  ||    |            |  ::::  ::||     :       |||:|||      :||         :    :
Db       1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127

Qy       1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
               |||              |      |      |      ||  ||:  ||      :      |:        |  ||  |     |         |::               |
Db       1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186

Qy       1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
               :: :        ::|:     | ::|:  :|   ||:  :|            |||    :                 |::|:::|:|:
Db       1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245

Qy       1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------- 1222
               |||:|||||  |||  |||::  |        |  ||              ||    ||  :      :      |
Db       1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
                      |    ::        |  |        |  ::  ||:                 :  :|  |      :  |:  :      |
Db       1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353

Qy       1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
```

```
                 :    |    :  | |   |      | ||:::| ::|||:       : :    |    |  : |
Db        1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413

Qy        1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
                 : :|||  |:     :     :::   |  : :  |  |  : ||||     |:   ||:
Db        1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465


RESULT 3
A2MG_RAT
ID   A2MG_RAT       STANDARD;      PRT;  1472 AA.
AC   P06238;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha-2-macroglobulin precursor (Alpha-2-M).
GN   A2M.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=87083488; PubMed=2432068;
RA   Gehring M.R., Shiels B.R., Northemann W., de Bruijn M.H.L.,
RA   Kan C.-C., Chain A.C., Noonan D.J., Fey G.H.;
RT   "Sequence of rat liver alpha 2-macroglobulin and acute phase control
RT   of its messenger RNA.";
RL   J. Biol. Chem. 262:446-454(1987).
RN   [2]
RP   SEQUENCE OF 1-164 FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=89160243; PubMed=2466233;
RA   Kunz D., Zimmermann R., Heisig M., Heinrich P.C.;
RT   "Identification of the promoter sequences involved in the
RT   interleukin-6 dependent expression of the rat alpha 2-macroglobulin
RT   gene.";
RL   Nucleic Acids Res. 17:1121-1138(1989).
RN   [3]
RP   SEQUENCE OF 178-227 AND 420-526 FROM N.A.
RX   MEDLINE=86033908; PubMed=2414291;
RA   Hayashida K., Okubo H., Noguchi M., Yoshida H., Kangawa K., Matsuo H.,
RA   Sakaki Y.;
RT   "Molecular cloning of DNA complementary to rat alpha 2-macroglobulin
RT   mRNA.";
RL   J. Biol. Chem. 260:14224-14229(1985).
CC   -!- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A
CC       UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,
CC       CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES
CC       FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT
CC       REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH
CC       TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST
CC       LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR
CC       WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE
CC       BAIT REGION A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE
CC       COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.
CC   -!- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-
CC       LINKED CHAINS.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- INDUCTION: BY INFLAMMATORY STIMULUS THE LEVEL OF THIS PROTEIN
CC       FIRST INCREASES, THEN DECREASES AFTER A MAXIMUM.
CC   -!- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC       INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J02635; AAA40636.1; -.
DR   EMBL; X13983; CAA32164.1; -.
DR   EMBL; X13984; CAA32164.1; JOINED.
DR   EMBL; X13985; CAA32164.1; JOINED.
DR   EMBL; M11792; AAA40637.1; -.
DR   EMBL; M11793; AAA40638.1; -.
DR   PIR; A26122; A26122.
DR   HSSP; P01023; 1BV8.
DR   GO; GO:0016975; F:alpha-2 macroglobulin; NAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW   Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
KW   Thioester bond.
FT   SIGNAL        1     27
FT   CHAIN        28   1472       ALPHA-2-MACROGLOBULIN.
FT   DOMAIN      620    750       BAIT REGION (APPROXIMATELY).
FT   DISULFID     52     90       BY SIMILARITY.
FT   DISULFID    254    302       BY SIMILARITY.
FT   DISULFID    272    290       BY SIMILARITY.
FT   DISULFID    281    434       BY SIMILARITY.
FT   DISULFID    473    473       INTERCHAIN (WITH OTHER CHAIN).
FT   DISULFID    598    769       BY SIMILARITY.
FT   DISULFID    647    694       BY SIMILARITY.
FT   DISULFID    819    847       BY SIMILARITY.
FT   DISULFID    845    881       BY SIMILARITY.
FT   DISULFID    919   1319       BY SIMILARITY.
FT   DISULFID   1077   1125       BY SIMILARITY.
FT   DISULFID   1350   1465       BY SIMILARITY.
FT   CROSSLNK    970    973       Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD     59     59       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     74     74       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    250    250       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    399    399       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    651    651       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    772    772       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    867    867       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    989    989       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1364   1364       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1422   1422       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1426   1426       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     23     23       L -> V (IN REF. 2).
FT   CONFLICT    490    490       M -> L (IN REF. 3).
SQ   SEQUENCE   1472 AA;  163701 MW;  D08FA99AEB10FB0B CRC64;

  Query Match            19.2%;  Score 1412.5;  DB 1;  Length 1472;
  Best Local Similarity  27.6%;  Pred. No. 1.8e-69;
  Matches  426;  Conservative  275;  Mismatches  580;  Indels  265;  Gaps   51;

Qy          5 PLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLK 64
              |||    || : | |  |     | ::|  | ::     |         |   |    | |:  |
Db         13 PLLLRLLLLLLPTDASA-PQKPIYMVMVPSLLH-AGTPEKACFLFSHLNETVAVRVSLES 70

Qy         65 TASN--LTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEI--YELRVTGRTQDEILFSNST 120
                 |  |    ::  : :|    ||         ::| :|:||:  : ::| | |  :    |   :
Db         71 VHGNQSLFTDLVVDKDLFHCTSF------TVPQSSSDELMFFTVQVKGATHE---FRRQS 121

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQW 179
               : : |     || |||| :|||  |:|:|:|     | |     : :|  |:|||:| | ||
Db        122 TVLVKKKESLVFAQTDKPIYKPGQTVRFRVVSLDESFHPLNELIPLLYIQDPKNNRIAQW 181

Qy        180 LSQQSDLGVISKTFQLSSHPILGDWSIQVQV-NDQTYYQSFQVSEYVLPKFEVTLQTPLY 238
               :    : |:    :| |||  |  | : : ::   : :|     | | |:|||||| :  |
Db        182 QNFNLEGGLKQLSFPLSSEPTQGSYKVVIRTESGRTVEHPFSVEEFVLPKFEVRVTVPET 241

Qy        239 CSMNSKHLNGTITAKYTYGKPVKGDVTLTFL-----PLSFWGKKKNI------------ 280
               ::  : :| ::    |||||| | ||:             | : :|::
Db        242 ITILEEEMNVSVCGIYTYGKPVPGRVTVNICRKYSNPSNCFGEESVAFCEKLSQQLDGRG 301

Qy        281 -------TKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTG 333
                     ||:|::     :     :     : :    |: |      :    |      || :| ::|
```

```
Db        302 CFSQLVKTKSFQLK-RQEYEMQLDVHAKIQEEGTGVEE----TGKG----LTKITRTITK 352

Qy        334 ISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITV 393
              :|                    | :  :  :  :  |    | :     |  :
Db        353 LS-----------------FVNVDSHFRQGIPFVGQVLLVDGRGTPIP----------- 383

Qy        394 TQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILED---SSELQLKAYF--- 447
                  |   : |::       ||     ||  |   :  |   :  |  |  |     :  | ::| :
Db        384 ----YETIFIGAD------EANLYINTTTDKHGLAR--FSINTDDIMGTSLTVRAKYKDS 431

Qy        448 --------LGSKSSMAVHSLFK--SPSKTYIQLKTRDENIKVGSPFEL----VVSGN--K 491
                      |  ::   |  |: :    |||::::  |::   :  ::       |:     :::|    :
Db        432 NACYGFRWLTEENVEAWHTAYAVFSPSRSFLHLESLPDKLRCDQTLEVQAHYILNGEAMQ 491

Qy        492 RLKEL--SYMVVSRGQLVAVG------KQNSTM--FS--LTPENSWTPKACVIVYYIEDD 539
              ||||  |:::::| :|  |        ||        ||  ::  |     |  | :::|  |  :
Db        492 ELKELVFYYLMMAKGGIVRAGTHVLPLKQGQMRGHFSILISMETDLAPVARLVLYAILPN 551

Qy        540 GEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLM 599
              ||::  |   |   ::    ||:  |  :           |:  :   |  :  :   |  |:  |:  |||:||  ||
Db        552 GEVVGDTAKYEIENCLANKVDLVFRPNSGLPATRALLSV-MASPQSLCGLRAVDQSVLLM 610

Qy        600 NASNDITMENV-------------------------VHELELYNTGYYLGMFMNS---- 629
                :::   :                           |  :  :  |   |        |:
Db        611 KPETELSASLIYDLLPVKDLTGFPQGADQREEDTNGCVKQNDTYINGILYSPVQNTNEED 670

Qy        630 -FAVFQECGLWVLTDANLTKDYI-DGVYDNAEYAERFMEENEGHIVDIHDFSLGSSPHVR 687
                 ::  ||  |  |::|:  |   :  :  :   ||            ::  |:       :  |     :    |
Db        671 MYGFLKDMGLKVFTNSNIRKPKVCERLRDNKGIPAAYHLVSQSHMDAFLESSESPTETRR 730

Qy        688 KHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQP 747
               :||||||          :       | ||||||:||  |  |   |  :|  |  ||||:     |:  ||||
Db        731 SYFPETWIWDLVVVDSAGVAEVEVTVPDTITEWKAGAFCLSNDTGLGLSPV-VQFQAFQP 789

Qy        748 FFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQT 807
              ||:  | :||||||||  |  |:  |:  |||           |  |  :|  |     |    |       |           |
Db        790 FFVELTMPYSVIRGEAFTLKATVLNYLPTCIRVAVQLEASP--DFLAAPEEKEQRSH--- 844

Qy        808 LLVPSEDGATVLFPIRPTHLGEIPITVTALSPTA---------------SDAVTQMILV 851
                :           |  :  :  |   ||  :   ||:|  :   :                |  :  :  :||
Db        845 -CICMNQRHTASWAVIPKSLGNVNFTVSAEALNSKELCGNEVPVVPEQGKKDTIIKSLLV 903

Qy        852 KAEGIEKSYSQSILLDLTDNRLQSTL-----KTLSFSFPPNTVTGSERVQITAIGDVLGP 906
               :  ||:|              ::||  :|  |    |     :  |     |   |    :  :|  :  :||:|
Db        904 EPEGLEN--------EVTFNSLLCPMGAEVSELIALKLPSDVVEESARASVTVLGDILGS 955

Qy        907 SINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELL 966
               ::         |::||||||||||:  |||||:||||  :  :|||   :|  ||::::  ||||:|
Db        956 AMQNTQDLLKMPYGCGEQNMVLFAPNIYVLDYLNETQQLTQEIKTKAIAYLNTGYQRQLN 1015

Qy        967 YQREDGSFSAFGNYDPSG----STWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQK 1022
               |:   |||:||||    |    |     :|||:||||:  |  :|   ||  ||:   :  :   ||    ||
Db       1016 YKHRDGSYSAFG--DKPGRNHANTWLTAFVLKSFAQARKYIFIDEVHITQALLWLSQQQK 1073

Qy       1023 SNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEF--- 1079
                ||  |   ·  |  ::::  ::||  :   |||:|||   :||         :    |:  ::·  |:: :
Db       1074 DNGCFRSSGSLLNNAMKGGVEDEVTLSAYITIALLEMSLPVTHPVVRNALFCLDTAWKSA 1133

Qy       1080 --SRGISDNYTLALITYALSSVGSPKA-KEALNMLTWRAEQEGGMQFWVSSESKSLSDS-- 1134
                 |  |    ||  ||:  ||  :   |:          |||    |     |     |:|        :   |  |  |
Db       1134 RGGAGGSHVYTKALLAYAFALAGNQDTKKEILKSLDEEAVKEEDSVHWTRPQ-KPSVSVG 1192

Qy       1135 --WQPR--SLDIEVAAYALLSHFLQFQTSEGIP----------IMRWLSRQRNSLGGFAS 1180
                :||:   :  :|:    ||  ||::    |||:                 |    |:|  :  |::||::||::||  |||:|
Db       1193 LWYQPQATSAEVEMTAYVLLAYL----TTEPAPTQEDLTAAMLIVKWLTKQQNSHGGFSS 1248

Qy       1181 TQDTTVALKALSEP-AALMNTERTNIQVTVTGPSSPSPLAVVQ---------------PM 1224
              ||||  |||  |||::  :         :      |||:        :    |       |                 |
Db       1249 TQDTVVALHALSKYGSATFTRAKKAAQVTIRSSGTFSTKFQVNNNNQLLLQRVTLPTVPG 1308

Qy       1225 AVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK----ENKDDLNHV 1280
                 :   |  |     |  ::  |:|                :  :|  |    |  |:     :|    |
Db       1309 DYTVKVTGEGCVYLQTSLKYSV-----------LPREEEFPFAVVVQTLPGTCEDPKAHT 1357

Qy       1281 DLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISLSET---VKKVEYDHGKLNLYL 1334
                :  :  |  |     |  ||:  :|  ::||||:        :  :  |       |  :  |    :    : :||
Db       1358 SFQISLNISYTGSRSESNMAIADVKMVSGFIPLKPTVKMLERSVHVSRTEVSNNHVLIYL 1417

Qy       1335 DSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
               ||  |:     ::       ::   :  :  :  |  |  :  ||||      ||    |::
Db       1418 DKVSNQTVNLSFTVQQDIPIRDLKPAVVKVYDYYEKDEFAVAKYSA 1463
```

```
RESULT 4
A2M1_MOUSE
ID   A2M1_MOUSE     STANDARD;      PRT;  1476 AA.
AC   P28665;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Murinoglobulin 1 precursor (MuG1).
GN   MUG1 OR MUG-1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 28-57.
RC   TISSUE=Liver;
RX   MEDLINE=91358495; PubMed=1840592;
RA   Overbergh L., Torrekens S., van Leuven F., van den Berghe H.;
RT   "Molecular characterization of the murinoglobulins.";
RL   J. Biol. Chem. 266:16903-16910(1991).
CC   -!- FUNCTION: A PROTEINASE ACTIVATES THE INHIBITOR BY SPECIFIC
CC       PROTEOLYSIS IN THE BAIT REGION, WHICH, BY AN UNKNOWN MECHANISM
CC       LEADS TO REACTION AT THE CYSTEINYL-GLUTAMYL INTERNAL THIOL ESTER
CC       SITE AND TO A CONFORMATIONAL CHANGE, WHEREBY THE PROTEINASE IS
CC       TRAPPED AND/OR COVALENTLY BOUND TO THE INHIBITOR. WHILE IN THE
CC       TETRAMERIC PROTEINASE INHIBITORS STERIC INHIBITION IS SUFFICIENTLY
CC       STRONG, MONOMERIC FORMS NEED A COVALENT LINKAGE BETWEEN THE
CC       ACTIVATED GLUTAMYL RESIDUE OF THE ORIGINAL THIOL ESTER AND A
CC       TERMINAL AMINO GROUP OF A LYSINE OR ANOTHER NUCLEOPHILIC GROUP ON
CC       THE PROTEINASE, FOR INHIBITION TO BE EFFECTIVE.
CC   -!- SUBUNIT: Monomer.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC       INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M65736; AAA73048.1; -.
DR   HSSP; P01023; 1BV8.
DR   MGD; MGI:99837; Mug1.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW   Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
KW   Multigene family; Thioester bond.
FT   SIGNAL        1     27
FT   CHAIN        28   1476       MURINOGLOBULIN 1.
FT   DOMAIN      677    734       BAIT REGION (APPROXIMATELY).
FT   DISULFID     48     86       BY SIMILARITY.
FT   DISULFID    251    276       BY SIMILARITY.
FT   DISULFID    269    288       BY SIMILARITY.
FT   DISULFID    461    555       BY SIMILARITY.
FT   DISULFID    587    773       BY SIMILARITY.
FT   DISULFID    634    680       BY SIMILARITY.
FT   DISULFID    849    885       BY SIMILARITY.
```

```
FT   DISULFID    923   1323       BY SIMILARITY.
FT   DISULFID   1081   1129       BY SIMILARITY.
FT   DISULFID   1354   1469       BY SIMILARITY.
FT   CROSSLNK    974    977       Isoglutamyl cysteine thioester (Cys-Gln)
FT                                (By similarity).
FT   CARBOHYD     55     55       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    313    313       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    500    500       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    749    749       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    776    776       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    871    871       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    993    993       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1142   1142       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1180   1180       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1426   1426       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1476 AA;  165139 MW;  D145D325B2DDDEAB CRC64;

  Query Match             19.1%;  Score 1406.5;  DB 1;  Length 1476;
  Best Local Similarity   27.7%;  Pred. No. 3.8e-69;
  Matches  435;  Conservative  270;  Mismatches  553;  Indels  311;  Gaps   50;

Qy         13 LCVCTAALA-------VAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKT 65
               ||: :  ||        :     :::|  |           :|    | :: :    |
Db          9 LCLFSVLLAFLHSASLLNGDSKYMVLVPS-----------QLYTETPEKICLHLYQLNE 56

Qy         66 ASNLTVSVLEAEG--------VFEKGSFKTLT--LPSLPLNSADEIYELRVTGRTQDEIL 115
                :| |:: |          | :|  |: ::  :|  |   :  ::    : |  |  :
Db         57 TVTVTASLVSQSGRKNLFDELVLDKDLFQCVSFIIPRLSSSDEEDFLYVDIKGPTHE--- 113

Qy        116 FSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNI-LIKDPKSN 174
              ||      :   :   |  ||:||||  :|||  |||||:::    :|      |  |:|||
Db        114 FSKRKAVLVKNKESVVFVQTDKPVYKPGQSVKFRVVSMDKMLRPLNELLPLAYIEDPKKN 173

Qy        175 LIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQV-QVNDQTYYQSFQVSEYVLPKFEVTL 233
               | ||    ::: |:    :| |::  ||  :  |  |  :  :  :      ||  | |:|||:|  | |
Db        174 RIMQWRDIKTENGLKQMSFSLAAEPIQGPYKIVVHKESGEKEEHSFTVMEFVLPRFNVDL 233

Qy        234 QTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFS 293
               : |    |:|: |: |:  |  |||||||| | |                     |||
Db        234 KVPNAMSVNDEVLSVTACGKYTYGKPVPGHV-------------------KINVCRETE 273

Qy        294 FNDEEMKNVMDSSNGLS----EYLDLSSPGP------VEILTTVTESVTGI--SRNVSTN 341
                  |: : :|  :|| |            :| |           :   ||||  ||:   ||: :|
Db        274 TGCREVNSQLD-NNGCSTQEVNITELQSKKRNYEVQLFHVNATVTEEGTGLEFSRSGTTK 332

Qy        342 VFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEY 401
                  :    |     :   :   :       |: :   :   |:  :|    ||    :  |
Db        333 IERITNKLI--FLKADSHFRHGIPFFVKVRLVDIKGDPIPNEK----VFIKAQELSYTSA 386

Qy        402 WSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKA-----------YFLGS 450
                 :    |                     | |:   :  || | :|            | :
Db        387 TTTDQHG----------------LAEFSIDTTCISGSS-LHIKVNHKEEDSCSYFYCMEE 429

Qy        451 KSSMAVHSLFK--SPSKTYIQLKTRDENI-------KVGSPFFELVVSGNKR-LKEL--SY 498
               : :  | |   :    | ||:|| | |    :|           | :|    :: |:  ||||   |
Db        430 RHASAKHVAYAVYSLSKSYIYLDTETSSILPCNQIHTVQAHF--ILKGDLGVLKELIFYY 487

Qy        499 MVVSRGQLVAVGKQNSTM------------FSLTPENSWTPKACVIVYYIEDDGEIISDV 546
               :|:::|  ::   |       :                :  |  |    |  | :::|  |     |||:|:|
Db        488 LVMAQGSIIQTGNHTHQVEPGEAPVKGKFALEIPVEFSMVPMAKMLIYTILPDGEVIADS 547

Qy        547 LKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDIT 606
               :    ::   :||: |:  |:  ::  |: | |:::  |:  | |:  | | |||:|||  |:   ::::
Db        548 VNFEIEKCLRNKVDLRFSTSQSLPASQTRLQVTAS-PQSLCGLRAVDQSVLLLKPESELS 606

Qy        607 MENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAE---- 662
                      :||    ||  |  |                  :  :  |::|  |   :  ::   ||
Db        607 PS------WIYNLP---GMQQNKF----------VPSSRLSEDQEDCILYSSWLAEKHTN 647

Qy        663 -----------RFMEEN-----------------EGHIVDIH----DFSLGSSPH----- 685
                         |::|:                   :  :|  |    :|  |  :|
Db        648 LVPHGTEKDVYRYVEDMGLTAFTNLMIKLPIICFDYGMVPISAPRVEFDLAFTPEISWSL 707

Qy        686 ------------------------VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
                                      :||:||||||:|    :        | |:||||:|| |
Db        708 RTTLSKRPEEPPRKDPSSNDPLTETIRKYFPETWVWDIVTVNSTGLAEVEMTVPDTITEW 767

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
                 |    :|  |  ||||::  |  ||||:|||:  ::|||||:|||  |  |:  |:  |||    : ::
Db        768 KAGALCLSNDTGLGLSSV-VPLQAFKPFFVEVSLPYSVVRGEAFMLKATVMNYLPTSMQM 826

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
               | :|  |    |  :    :        |: : :         |   : : |  ||  :  :|:| :
Db        827 SVQLEASPDFTAVPVGDD------QDSYCLSANGRHTSSWLVTPKSLGNVNFSVSAEAQQ 880

Qy        841 AS----------------DAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSF 884
               :|             | |  ::::|:  |||::  ::  | |    :|   :    : :|
Db        881 SSEPCGSEVATVPATGRKDTVVKVLIVEPEGIKQEHTFSSLFCASDAEIS---EKMSSGP 937

Qy        885 PPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQ 944
               ||   |    |  |    :  :||:|  :|         :   :  |||||||||:  ||||||:|  || :  :|
Db        938 PPTVVKDSARAHFSVMGDILSSAIRNTQNPLHMPYGCGEQNMVLFAPNIYVLKYLNETQQ 997

Qy        945 LTDNLKEKALSFMRQGYQRELLYQREDGSFSAFG--NYDPSGSTWLSAFVLRCFLEADPY 1002
               ||   :|  |||  |:|  ||||||  |: :|||:||||   |  :   |:|||:||||:  |  :|   :
Db        998 LTQKIKTKALGFLRAGYQRELNYKHKDGSYSAFGDQNGEREGNTWLTAFVLKSFAQARAF 1057

Qy       1003 IDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKY 1062
               | ||:: :   .:|||     || ||  |     |  : :: ::||     :||:|||   :||
Db       1058 IFIDESHITHAFTWLSQKQKDNGCFRSSGSLFNNAMKGGVDDEMTLSAYITMALLESSLP 1117

Qy       1063 QPNIDVQESIHFLES-----EFSRGISDNYTLALITYALSSVGSP-KAKEALNMLTWRAE 1116
               :  | :::  |||      |   |   |   ||  ||:  ||  :   |:   |   |   |   |
Db       1118 ATHPVVSKALSCLESSWKTIEQERNASFVYTKALMAYAFALAGNQNKRDEILKSLDEEAI 1177

Qy       1117 QEGGMQFWV-SSESKSLSDSW----QPRSLDIEVAAYALLSHFLQFQTSEGIP-------- 1163
               :|           :|: |:            |  ::|: || :|:             |:: |
Db       1178 KENNSIHWKRPQKSRKSEHHLYKPQASSAEVEMNAYVVLARL----TAQPAPSPEDLTLS 1233

Qy       1164 --IMRWLSRQRNSLGGFASTQDTTVALKALSEFAAL---MNTERTNIQVTVTGPSSP--- 1215
                  : ||::|:|| |||:||||| ||| |||::  |:        :   :  |  :   ||   |
Db       1234 MSTIMWLTKQQNSNGGFSSTQDTVVALDALSKYGAVTFSRSQKTTLVTIQSTGSFSQKFQ 1293

Qy       1216 ---SPLAVVQPMAV-------NISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFD 1265
                 |    ::|  :|:                  || :|  |     |   :  ||:                :    ||
Db       1294 VENSNRLLLQQVALPDIPGDYTISVSGEGCVYAQTMLRYNMHLE---------KQLSAFA 1344

Qy       1266 LDVAVK----ENKDDLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGFMVPSEAISLSETV 1320
               :  |          |     |     :::  |::|   |   | :  :| :|||     |   |   ||
Db       1345 IWVQTVPLTCNNPKGHNSFQISLEISYTGSRPASNMVIADVKMLSGF------IPLKPTV 1398

Qy       1321 KKV---------EYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPR 1371
               ||:             |   :   :|||  |        :     ::   |  |  |  |  |  |  :|||
Db       1399 KKLERLEHVSRTEVSNNNVLIYLDQVTNQTLAFSFIIQQDIPVRNLQPAIVKVYDYYETD 1458

Qy       1372 RQAVRSYNS 1380
                  |   |:|
Db       1459 EMAFAEYSS 1467

RESULT 5
A2M2_MOUSE
ID   A2M2_MOUSE     STANDARD;      PRT;  1451 AA.
AC   P28666;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Murinoglobulin 2 precursor (MuG2).
GN   MUG2 OR MUG-2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Liver;
RX   MEDLINE=91358495; PubMed=1840592;
RA   Overbergh L., Torrekens S., van Leuven F., van den Berghe H.;
RT   "Molecular characterization of the murinoglobulins.";
RL   J. Biol. Chem. 266:16903-16910(1991).
CC   -!- FUNCTION: A PROTEINASE ACTIVATES THE INHIBITOR BY SPECIFIC
CC       PROTEOLYSIS IN THE BAIT REGION, WHICH, BY AN UNKNOWN MECHANISM
CC       LEADS TO REACTION AT THE CYSTEINYL-GLUTAMYL INTERNAL THIOL ESTER
CC       SITE AND TO A CONFORMATIONAL CHANGE, WHEREBY THE PROTEINASE IS
CC       TRAPPED AND/OR COVALENTLY BOUND TO THE INHIBITOR. WHILE IN THE
CC       TETRAMERIC PROTEINASE INHIBITORS STERIC INHIBITION IS SUFFICIENTLY
CC       STRONG, MONOMERIC FORMS NEED A COVALENT LINKAGE BETWEEN THE
CC       ACTIVATED GLUTAMYL RESIDUE OF THE ORIGINAL THIOL ESTER AND A
CC       TERMINAL AMINO GROUP OF A LYSINE OR ANOTHER NUCLEOPHILIC GROUP ON
CC       THE PROTEINASE, FOR INHIBITION TO BE EFFECTIVE.
CC   -!- SUBUNIT: Monomer.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC       INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M65238; AAA73041.1; -.
DR   PIR; B41185; B41185.
DR   HSSP; P01023; 1BV8.
DR   MGD; MGI:99836; Mug2.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW   Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
KW   Multigene family; Thioester bond.
FT   SIGNAL        1     27       BY SIMILARITY.
FT   CHAIN        28   1451       MURINOGLOBULIN 2.
FT   DOMAIN      677    734       BAIT REGION (APPROXIMATELY).
FT   DISULFID     48     86       BY SIMILARITY.
FT   DISULFID    251    276       BY SIMILARITY.
FT   DISULFID    269    288       BY SIMILARITY.
FT   DISULFID    461    555       BY SIMILARITY.
FT   DISULFID    587    773       BY SIMILARITY.
FT   DISULFID    634    680       BY SIMILARITY.
FT   DISULFID    849    885       BY SIMILARITY.
FT   DISULFID    923   1274       BY SIMILARITY.
FT   DISULFID   1081   1104       BY SIMILARITY.
FT   DISULFID   1298   1444       BY SIMILARITY.
FT   CROSSLNK    974    977       Isoglutamyl cysteine thioester (Cys-Gln)
FT                                (By similarity).
FT   CARBOHYD     55     55       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    313    313       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    500    500       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    749    749       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    776    776       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    871    871       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1401   1401       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1451 AA;  162368 MW;  9B2A82EFAA6258F6 CRC64;

  Query Match            19.1%;  Score 1402;  DB 1;  Length 1451;
  Best Local Similarity  27.7%;  Pred. No. 6.6e-69;
  Matches  439;  Conservative  250;  Mismatches  527;  Indels  370;  Gaps   54;

Qy         13 LCVCTAALAVAPG-------PRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKT 65
                ||: :  ||  |             :::|  |               :|   | :: :     |
Db          9 LCLFSVLLAFLPSASSLNGDSKYMVLVPS------------QLYTETPEKICLHLYHLNE 56

Qy         66 ASNLTVSVLEAEG--------VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEI 114
                 :| |::    |           | :|  |: ::      ||| ||    :| : : | | :
Db         57 TVTVTASLVSQTGRRNLFDELVVDKDLFQCVSFIIPTLNSPDEEEFLY-VDIKGPTHE-- 113

Qy        115 LFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNI-LIKDPKS 173
               ||     :  : |    ||:||||  :|||  |  ||||:|::      :|       | :   |:|||
Db        114 -FSKRNAVLVKNKESVVFVQTDKPVYKPGQSVKFRVVSMDKTLRPLNELLPLAYIEDPKK 172

Qy        174 NLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQ----VNDQTYYQSFQVSEYVLPKF 229
               | |  ||     ::: |:     :|  |::  ||  |  : |  |      | ::       || |  |:|||: |
Db        173 NRIMQWRDIKTENGLKQMSFSLAAEPIQGPYKIVVHKQSGVKEE---HSFTVMEFVLPRF 229

Qy        230 EVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGS 289
                |  | :  |    |:| :  |   |:  ||||||||  |  |  ::        :  |  |  : :        ||
Db        230 NVDLKVPNAISVNDEVLQVTVCGKYTYGKPVPGQVKISICHETEAGCKE-VNSKLDNNGC 288

Qy        290 ANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRN---------VST 340
               :     |   |::      |      :   :   |           :  ||||   ||:   |              ::
Db        289 STQEVNITELQ-----SKKRNYEVQL-----FHVNATVTEEGTGLEFNGYGTTKIERITN 338

Qy        341 NVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTE 400
                : |  :  |              :   :   :   |    |::       |: :   |      |  |      ||
Db        339 KLIFLKAD---------SHFRHGIPFFVKVRLVDIKGDPIPNE----RVFIKAQVLGYTS 385

Qy        401 YWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKS-------- 452
                  :  |    |          :::  :|              ::  :|       |  :|    |   |
Db        386 TTTTDQHGLAK--------FSIDTAGF---------SGSSLHIKVNHKGKDSCYFFYCME 428

Qy        453 ---------SMAVHSLFKSPSKTYIQLKTRDENI-------KVGSPFFELVVSGNKR-LKE 495
                          :  ||:||     ||:||  |        :|            |  :  |   ::  |:    |||
Db        429 ERYASAEHVAYAVYSL----SKSYIYLVKETSSILPCNQIHTVQAHF--ILKGDLGVLKE 482

Qy        496 L--SYMVVSRGQLVAVGKQNSTM-----------FSLTPENSWTPKACVIVYYIEDDGE 541
               |     |:|:::|  ::   |           :                :   |   |   |   | :::|   |    |||
Db        483 LVFYYLVMAQGSIIQTGNHTHQVEPGEAPVKGNFDLEIPVEFSMAPMAKMLIYTILPDGE 542

Qy        542 IISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNA 601
               :|:|  :    ::    :||:  |  :|   ::  |:  :   |:::  :   |  |:  |:  |||:||   |:
Db        543 VIADSVNFEIEKCLRNKVDLSFSSSQSLPASQTRLQVTAS-PQSLCGLRAVDQSVLLLKP 601

Qy        602 SNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYA 661
                :::: |        :||    ||   |             :   ::|::|   |  :   :
Db        602 EDELSPS------WIYNLP---GMQHNKF----------IPSSSLSEDREDCILYSSWVA 642

Qy        662 E--------------RFMEENEGHIVDIHDFS-------------------------- 679
               |                  |::|:      :|:   |:
Db        643 EKHTDWVPHGREKDVYRYVED-----MDLKAFTNLKIKLPKICFDSAPMSGPRGKFDLAF 697

Qy        680 ---------LGSSPH-------------------VRKHFPETWIWLDTNMGYRIYQEFE 710
                          |||                          :||:|||||:|       :       |  |
Db        698 SSEVSGTLQKGSSKRPQPEEPPREDPPPKDPLAETIRKYFPETWVWDIVTVNSTGVAEVE 757

Qy        711 VTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITI 770
               :||||:||  |  |      :|  |  ||||::   |  |||||||||: ::|||||:|||  |  |: |:
Db        758 MTVPDTITEWKAGALCLSNDTGLGLSSV-VPLQAFQPFFVEVSLPYSVVRGEAFMLKATV 816

Qy        771 FNYLKDATEVKVIIEKSDKFDILMTSSE-----INATG-HQQTLLVPSEDGATVLFPIRP 824
               |||    :    :|   :|    |      :        ::|  |   |  ||
Db        817 MNYLPTSMRMSVQLEASPDFTAVPVGDDHDSYCLSANGRHTSSWLV-----------TP 864

Qy        825 THLGEIPITVTALSPTAS---------------DAVTQMILVKAEGIEKSYSQSILLDL 868
                 ||  :    :|:  :   :|                  |   |  ::::|:  |||:: :: : |
Db        865 KSLGNVNFSVSVEAQQSSEPCGSEVATVPATGRKDTVVKVLIVEPEGIKQEHTFNSLFCA 924

Qy        869 TDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMIN 928
               :|   :       : :|      ||    |    |    |    : :||:|  :|    :|:  |||||||||:
Db        925 SDAEIS---EKMSLVLPPTVVKDSARAHFSVMGDILSSAIKNTQNLLHMPYGCGEQNMVL 981

Qy        929 FAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFG--NYDPSGST 986
               ||||||:|  ||  |  :|||    :|  |||  |:|  ||||||  |: :|||:||||  ·  :   |:|
Db        982 FAPNIYVLKYLDKTQQLTQKIKTKALGFLRAGYQRELNYKHKDGSYSAFGDQNGEREGNT 1041

Qy        987 WLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPV 1046
               ||:|||:  |  :|   :|  ||::  :     :|||      ||  ||  |       |   :  |:::        |  ||
```

```
Db      1042 WLTAFVLKSFAQARAFIFIDESHITHAFTWLSQQQKDNGCFRSSGSLFHNDI----KHPV 1097

Qy      1047 TLTAYIVTSLLGYRKYQPNIDVQESIHFLES-----EFSRGISDNYTLALITYALSSVGS 1101
                                   | :::  |||      |  |   :  || ||: || :  |:
Db      1098 --------------------VSKALSCLESSWKTIEQGRNANFVYTKALMAYAFALAGN 1136

Qy      1102 -PKAKEALNMLTWRAEQEGGMQFWV------SSESKLSDSWQPRSLDIEVAAYALLSHFL 1154
                |  |  |  |    |  :|        |            ||  |     |  |:::|:  || :|:
Db      1137 QDKRNEILKSLDEEAIKEDNSIHWERPQKPRKSEHNLYKP-QASSIEVEMNAYVVLARL- 1194

Qy      1155 QFQTSEGIP---------IMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN 1204
                |::  |              :  ||::|:||  |||:|||||  |||  |||::  |:  :  |
Db      1195 ---TAQPAPSPEDLTLSRSTIMWLTKQQNSNGGFSSTQDTVVALDALSKYGAVTFSRRQK 1251

Qy      1205 IQ-VTVTGPSSPSPLAVVQ--------------PMAVNISANGFGFAICQLNVVYNVKA 1248
                 ||:      |  |      |:                |        ||  :|  |     |    :  ||:
Db      1252 TSLVTIQSTGSFSQKFQVENSNCLLLQQVPLPDIPGDYTISVSGEGCVYAQTTLRYNMHL 1311

Qy      1249 SGSSRRRRSIQNQEAFDLDVAVK----ENKDDLNHVDLNVCTSFSGP-GRSGMALMEVNL 1303
                          :  |  ||  |  |            |       |      :::    |::|       |  |  :  :|  :
Db      1312 E---------KQQSAFALRVQTVPLTCNNPKGHNSFQISLEISYTGSRPASNMVIADVKM 1362

Qy      1304 LSGFMVPSEAISLSETVKKV---------EYDHGKLNLYLDSVNETQFCVNIPAVRNFKV 1354
                ||||          |  |. ||||:              |    :    :  ||||  |            :         ::    |
Db      1363 LSGF------IPLKPTVKKLERLEHISRTEVSNNNVLLYLDQVTNQTLAFSFIIQQDISV 1416

Qy      1355 SNTQDASVSIVDYYEPRRQAVRSYNS 1380
                |  |  |  |  :  ||||      |      |:|
Db      1417 RNLQPAIVKVYDYYETDEVAYAEYSS 1442
```

```
RESULT 6
PZP_HUMAN
ID   PZP_HUMAN     STANDARD;      PRT;  1482 AA.
AC   P20742; Q15273;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Pregnancy zone protein precursor.
GN   PZP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91113734; PubMed=1989698;
RA   Devriendt K., van den Berghe H., Cassiman J.-J., Marynen P.;
RT   "Primary structure of pregnancy zone protein. Molecular cloning of a
RT   full-length PZP cDNA clone by the polymerase chain reaction.";
RL   Biochim. Biophys. Acta 1088:95-103(1991).
RN   [2]
RP   SEQUENCE OF 670-759.
RX   MEDLINE=89380162; PubMed=2476433;
RA   Sottrup-Jensen L., Sand O., Kristensen L., Fey G.H.;
RT   "The alpha-macroglobulin bait region. Sequence diversity and
RT   localization of cleavage sites for proteinases in five mammalian
RT   alpha-macroglobulins.";
RL   J. Biol. Chem. 264:15781-15789(1989).
RN   [3]
RP   SEQUENCE OF 668-753 FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=90242963; PubMed=1692292;
RA   Marynen P., Devriendt K., van den Berghe H., Cassiman J.-J.;
RT   "A genetic polymorphism in a functional domain of human pregnancy
RT   zone protein: the bait region. Genomic structure of the bait domains
RT   of human pregnancy zone protein and alpha 2 macroglobulin.";
RL   FEBS Lett. 262:349-352(1990).
RN   [4]
RP   SEQUENCE OF 974-983.
RX   MEDLINE=86059456; PubMed=2415522;
RA   Sand O., Folkersen J., Westergaard J.G., Sottrup-Jensen L.;
RT   "Characterization of human pregnancy zone protein. Comparison with
RT   human alpha 2-macroglobulin.";
RL   J. Biol. Chem. 260:15723-15735(1985).
RN   [5]
RP   SEQUENCE OF 1259-1461 FROM N.A.
RX   MEDLINE=90034205; PubMed=2478422;
RA   Devriendt K., Zhang J., van Leuven F., van den Berghe H.,
RA   Cassiman J.-J., Marynen P.;
RT   "A cluster of alpha 2-macroglobulin-related genes (alpha 2 M) on
RT   human chromosome 12p: cloning of the pregnancy-zone protein gene and
RT   an alpha 2M pseudogene.";
RL   Gene 81:325-334(1989).
RN   [6]
RP   INTERACTIONS WITH PROTEINASES AND METHYLAMINE.
RX   MEDLINE=90122795; PubMed=2692707;
RA   Christensen U., Simonsen M., Harrit N., Sottrup-Jensen L.;
RT   "Pregnancy zone protein, a proteinase-binding macroglobulin.
RT   Interactions with proteinases and methylamine.";
RL   Biochemistry 28:9324-9331(1989).
CC   -!- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A
CC       UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,
CC       CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES
CC       FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT
CC       REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH
CC       TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST
CC       LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR
CC       WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE
CC       BAIT REGION A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE
CC       COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.
CC   -!- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-
CC       LINKED CHAINS.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC       INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X54380; CAA38255.1; -.
DR   EMBL; M24416; AAA60234.1; -.
DR   EMBL; X51541; CAA35919.1; -.
DR   PIR; S13495; S13495.
DR   HSSP; P01023; 1BV8.
DR   Genew; HGNC:9750; PZP.
DR   MIM; 176420; -.
DR   GO; GO:0016975; F:alpha-2 macroglobulin; TAS.
DR   GO; GO:0007565; P:pregnancy; TAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW   Serine protease inhibitor; Glycoprotein; Bait region; Plasma; Signal;
KW   Thioester bond.
FT   SIGNAL        1     25
FT   CHAIN        26   1482       PREGNANCY ZONE PROTEIN.
FT   DOMAIN      685    735       BAIT REGION.
FT   CROSSLNK    978    981       Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD     24     24       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     54     54       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     69     69       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    392    392       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    406    406       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    753    753       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    875    875       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    932    932       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    997    997       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1430   1430       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CONFLICT    691    691       V -> M (IN REF. 3).
FT   CONFLICT    753    753       N -> Q (IN REF. 2).
SQ   SEQUENCE   1482 AA;  163836 MW;  01773EF61D2FCC3B CRC64;

  Query Match             19.0%;  Score 1396.5;  DB 1;  Length 1482;
  Best Local Similarity   28.3%;  Pred. No. 1.4e-68;
  Matches  432;  Conservative  280;  Mismatches  587;  Indels  225;  Gaps   53;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
Db         14 LLILLSASDSNSTEPQYMVLVPSLLHTEAPKK-GCVLLSHLNETVTVSASLESGRENRS- 71

Qy         72 SVLEAEGVFEKGSFK--TLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRI 129
Db         72 --LFTDLVAEKDLFHCVSFTLPRISASSEVAFLSIQIKGPTQD--FRKRNTVLVLNTQSL 127

Qy        130 SVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSDLGV 188
Db        128 -VFVQTDKPMYKPGQTVRFRVVSVDENFRPRNELIPLIYLENPRRNRIAQWQSLKLEAGI 186

Qy        189 ISKTFQLSSHPILGDWSIQVQVNDQTYYQ-SFQVSEYVLPKFEVTLQTPLYCSMNSKHLN 247
Db        187 NQLSFPLSSEPIQGSYRVVVQTESGGRIQHPFTVEEFVLPKFEVKVQVPKIISIMDEKVN 246

Qy        248 GTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSN 307
Db        247 ITVCGEYTYGKPVPGLAT-----VSLCRKLSRVLNCDKQEVCEEFS---QQL-----NSN 293

Qy        308 GL------SEYLDLSSPG---PVEILTTVTESVTGI---SRNVS--TNVFFKQHDYIIEF 353
Db        294 GCITQQVHTKMLQITNTGFEMKLRVEARIREEGTDLEVTANRISEITNIVSK-----LKF 348

Qy        354 FDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSG----- 408
Db        349 VKVDSHFRQGIPFFA--QVLLVDGKGVPIPNKL--FFISVNDANYYSNATTNEQGLAQFS 404

Qy        409 -NQKMEAVQKI---NYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPS 464
Db        405 INTTSISVNKLFVRVFTVHPNLCFHYSW-VAEDHQGAQ--------HTANRVFSL----S 451

Qy        465 KTYIQLKTRDENIKVG------SPFELVVSGNKRLKELS--YMVVSRGQLVAVG------ 510
Db        452 GSYIHLEPVAGTLPCGHTETITAHYTLNRQAMGELSELSFHYLIMAKGVIVRSGTHTLPV 511

Qy        511 ----KQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV 566
Db        512 ESGDMKGSFALSFPVESDVAPIARMFIFAILPDGEVVGDSEKFEIENCLANKVDLSFSPA 571

Qy        567 KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELE----------- 615
Db        572 QSPPASHAHLQVAAA-PQSLCALRAVDQSVLLMKPEAELSVSSVYNLLTVKDLTNFPDNV 630

Qy        616 ---------------LYNTGYYLGMFMNS---FAVFQECGLWVLTDANLTKDYIDGVYDN 657
Db        631 DQQEEEQGHCPRPFFIHNGAIYVPLSSNEADIYSFLKGMGLKVFTNSKIRKPKSCSVIPS 690

Qy        658 AE--------------YAERFMEENEG--HIVDI-HDFSLGSSPH-VRKHFPETWIWLDT 699
Db        691 VSAGAVGQGYYGAGLGVVERPYVPQLGTYNVIPLNNEQSSGPVPETVRSYFPETWIWELV 750

Qy        700 NMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVI 759
Db        751 AVNSSGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISST-ASLRAFQPFFVELTMPYSVI 809

Qy        760 RGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVL 819
Db        810 RGEVFTLKATVLNYLPKCIRVSVQLKASPAF-----LASQNTKGEESYCICGSE-RQTLS 863

Qy        820 FPIRPTHLGEIPITVTALSPTA----------------SDAVTQMILVKAEGIEKSYSQS 863
Db        864 WTVTPKTLGNVNFSVSAEAMQSLELCGNEVVEVPEIKRKDTVIKTLLVEAEGIEQEKTFS 923

Qy        864 ILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGE 923
Db        924 ---SMTCASGANVSEQLSLKLPSNVVKESARASFSVLGDILGSAMQNIQNLLQMPYGCGE 980

Qy        924 QNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFG-NYDP 982
Db        981 QNMVLFAPNIYVLNYLNETQQLTQEIKAKAVGYLITGYQRQLNYKHQDGSYSTFGERYGR 1040

Qy        983 S-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG 1041
Db       1041 NQGNTWLTAFVLKTFAQARSYIFIDEAHITQSLTWLSQMQKDNGCFRSSGSLLNNAIKGG 1100

Qy       1042 NKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLES------EFSRGISDNYTLALITYA 1095
Db       1101 VEDEATLSAYVTIALLEIPLPVTNPIVRNALFCLESAWNVAKEGTHG-SHVYTKALLAYA 1159

Qy       1096 LSSVGSP-KAKEALNMLTWRAEQEGGMQFWVSSESKLSD-----SWQPRSLDIEVAAYAL 1149
Db       1160 FSLLGKQNQNREILNSLDKEAVKEDNLVHWERPQRPKAPVGHLYQTQAPSAEVEMTSYVL 1219

Qy       1150 LSHFLQFQTSEGIP----------IMRWLSRQRNSLGGFASTQDTTVALKALSEF--AAL 1197
Db       1220 LAYL----TAQPAPTSGDLTSATNIVKWIMKQQNAQGGFSSTQDTVVALHALSRYGAATF 1275

Qy       1198 MNTERTNIQVTVTGPSSPSP--------LAVVQPMAVN-------ISANGFGFAICQLNV 1242
Db       1276 TRTEKT-AQVTVQDSQTFSTNFQVDNNNLLLLQQISLPELPGEYVITVTGERCVYLQTSM 1334

Qy       1243 VYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLN--HVDLNVCTSFSGPGR---SGMA 1297
Db       1335 KYNILPE---------KEDSPFALKVQTVPQTCDGHKAHTSFQISLTISYTGNRPASNMV 1385

Qy       1298 LMEVNLLSGFMVPSEAISL---SETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKV 1354
Db       1386 IVDVKMVSGFIPLKPTVKMLERSSSVSRTEVSNNHVLIYVEQVTNQTLSFSFMVLQDIPV 1445

Qy       1355 SNTQDASVSIVDYYEPRRQAVRSY 1378
Db       1446 GDLKPAIVKVYDYYETDESVVAEY 1469

RESULT 7
OVOS_CHICK
ID   OVOS_CHICK     STANDARD;      PRT;  1473 AA.
AC   P20740;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   Ovostatin precursor (Ovomacroglobulin).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Oviduct;
RX   MEDLINE=95218210; PubMed=7535598;
RA   Nielsen K.L., Sottrup-Jensen L., Nagase H., Thoegersen H.C.,
RA   Etzerodt M.;
RT   "Amino acid sequence of hen ovomacroglobulin (ovostatin) deduced from
RT   cloned cDNA.";
RL   DNA Seq. 5:111-119(1994).
RN   [2]
RP   SEQUENCE OF 37-49.
RC   TISSUE=Egg white;
RX   MEDLINE=83238315; PubMed=6408074;
RA   Nagase H., Harris E.D. Jr., Woessner J.F., Brew K.;
RT   "Ovostatin: a novel proteinase inhibitor from chicken egg white. I.
RT   Purification, physicochemical properties, and tissue distribution of
RT   ovostatin.";
RL   J. Biol. Chem. 258:7481-7489(1983).
RN   [3]
RP   SEQUENCE OF 976-1028.
```

```
RC   TISSUE=Egg white;
RX   MEDLINE=93192299; PubMed=7680577;
RA   Nielsen K.L., Sottrup-Jensen L.;
RT   "Evidence from sequence analysis that hen egg-white ovomacroglobulin
RT   (ovostatin) is devoid of an internal beta-Cys-gamma-Glu thiol
RT   ester.";
RL   Biochim. Biophys. Acta 1162:230-232(1993).
CC   -!- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A
CC       UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,
CC       CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES
CC       FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT
CC       REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH
CC       TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST
CC       LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR
CC       WEIGHT SUBSTRATES IS GREATLY REDUCED).
CC   -!- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-
CC       LINKED CHAINS.
CC   -!- PTM: THIS PROTEIN LACKS THE THIOESTER BOND FOUND IN OTHER MEMBERS
CC       OF THIS FAMILY.
CC   -!- PTM: CONTAINS 56 MOL GLUCOSAMINE PER MOL SUBUNIT.
CC   -!- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC       INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X78801; CAA55384.1; -.
DR   EMBL; X78801; CAA55385.1; ALT_INIT.
DR   PIR; I50671; A20872.
DR   HSSP; P01023; 1BV8.
DR   GO; GO:0017114; F:wide-spectrum protease inhibitor activity; NAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; FALSE_NEG.
KW   Serine protease inhibitor; Glycoprotein; Bait region; Signal.
FT   SIGNAL        1     36
FT   CHAIN        37   1473       OVOSTATIN.
FT   CARBOHYD     67     67       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     82     82       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     89     89       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    191    191       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    342    342       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    403    403       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    527    527       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    588    588       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    757    757       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1141   1141       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1221   1221       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1315   1315       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1347   1347       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     46     46       M -> V (IN REF. 2).
SQ   SEQUENCE   1473 AA;  166354 MW;  A33C6847A14179BF CRC64;

  Query Match             19.0%;  Score 1395;  DB 1;  Length 1473;
  Best Local Similarity   27.0%;  Pred. No. 1.6e-68;
  Matches  421;  Conservative  309;  Mismatches  522;  Indels  308;  Gaps   59;

Qy         18 AALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVS---VL 74
                ||     | |::::  | :::                |||| ::   |    | |:|   ||
Db         33 AAAGKEPEPQYVLMVPAVLQ------------SDSPSQVCLQFFNL----NQTISVRVVL 76

Qy         75 EAE----GVFEKGSFKTLTLPSL-----PLNSADEIY-ELRVTGRTQD-----EILFSNS 119
              |  :       :|||  :   :   |      |:  |      :    |  |  |       ::  |
Db         77 EYDTINTTIFEKNTTTSNGLQCLNFMIPPVTSVSLAFISFTAKGTTFDLKERRSVMIWNM 136

Qy        120 TRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQ 178
              ||        ||:||||| :|||  |  | ||:| |    :|||  :     :: ::||::| | |
Db        137 E--SF------VFVQTDKPIYKPGQSVMFRVVALDFNFKPVQEMYPLIAVQDPQNNRIFQ 188

Qy        179 WLSQQSDLGVISKTFQLSSHPILGDWSIQV-QVNDQTYYQSFQVSEYVLPKFEVTLQTPL 237
              | :   |:: ::    | |:  ||||::  |  | : : :        || |  |||||||:||:  |
Db        189 WQNVTSEINIVQIEFPLTEEPILGNYKIIVTKKSGERTSHSFLVEEYVLPKFDVTVTAPG 248

Qy        238 YCSMNSKHLNGTITAKYTYGKPVKGDVTLTFL-PLSFWGKKK------NITKTFKINGSA 290
                ::        |  |  ||||:||:|  |   |:          :|: |         :  ||    :|
Db        249 SLTVMDSELTVKICAVYTYGQPVEGKVQLSVCRDFDSYGRCKKSPVCQSFTKDLDTDGCL 308

Qy        291 NFSFNDE--EMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHD 348
              :   :  :  |:  :        |      ||         |       :|| |    |  |   ::    :   |
Db        309 SHILSSKVFELNRI-----GYKRNLD------VKAIVTEKEQVCNLTATQSISI--TQVM 355

Qy        349 YIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSG 408
                ::|  :       :   :  :   :|:    |  :  :         :|  ||  :    |    :
Db        356 SSLQFENVDHHYRRGIPYFGQIKLVDKDNSPI------SNKVIQLFVNNKNTH------- 402

Qy        409 NQKMEAVQKINYTVPQSG--TFKIEFPILEDSSELQLKAYFLGS-------------KS 452
                                 |:|   :|   | |:   :  |  || ||| :  |                         :
Db        403 ----------NFTTDINGIAPFSIDTSKIFD-PELSLKALYKTSDQCHSEGWIEPSYPDA 451

Qy        453 SMAVHSLFKSPSKTYIQLKTRDENIKVGSP------FELVVSGNKRLKELS--YMVVSRG 504
              |::|  |:  | : :::::   :::  |                :  |     |  :  :   :  |: :::|
Db        452 SLSVQRLY-SWTSSFVRIEPLWKDMSCGQKRMITVYYILNTEGYEHINIVNFYYVGMAKG 510

Qy        505 QLVAVGK--------QNST-MFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVF 555
              ::|   |:            || |  |   |          |     :|| |  :        |::|  :   :   |
Db        511 KIVLTGEIKVNIQADQNGTFMIPLVVNEKMAPALRLLVYMLHPAKELVADSVRFSIEKCF 570

Qy        556 KNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENV--VHE 613
              |||::|  :|:  :    :   |||  |       :|     :  |||||:  |:  :   :::  |  :  :|
Db        571 KNKVQLQFSEKQMLTTSNVSLVIEAA-ANSFCAVRAVDKSMLLLKSETELSAETIYNLHP 629

Qy        614 L------------------------ELYNTGYYL-----GMFMNSFAVFQECGLWVLTD 643
              :                        :::: |  |         |:  :  :    ::  |:    |:
Db        630 IQDLQGYIFNGLNLEDDPQDPCVSSDDIFHKGLYYRPLTSGLGPDVYQFLRDMGMKFFTN 689

Qy        644 ANL------TKDYI--------------DGVYDNAEYAERFMEENEGHIVDIHDFSLGS 682
              : :    |:: :                       |  :||  |       |  :  ||::
Db        690 SKIRQPTVCTRETVRPPSYFLNAGFTASTHHVKLSAEVARE--ERGKRHILET------- 740

Qy        683 SPHVRKHFPETWIW---LDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTP 739
                  :|: ||||||      |   :   |       |:||:|   :|||  :  |  :  |   |   :  |
Db        741 ---IREFFPETWIWDIILINSTG---KASVSYTIPDTITEWKASAFCVEELAGFGM-SVP 793

Qy        740 VELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEI 799
                | ||||||: | ||||:| ||:| :    :||||      :: |::
Db        794 ATLTAFQPFFVDLTLPYSIIHGEDFLVRANVFNYLNHCIKINVLL--------------L 839

Qy        800 NATGHQQTLLVPSEDGA-------TVLFPIRPTHLGEIPITVTA---------------L 837
                :  :|  |: | :||       : ::  |  |   |::  ::||                                :
Db        840 ESLDYQAKLISPEDDGCVCAKIRKSYVWNIFPKGTGDVLFSITAETNDDEACEEEALRNI 899

Qy        838 SPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQI 897
                    |    : :||:  |||    :|: |: : |: :    : ::    |  |  |  ||  |
Db        900 RIDYRDTQIRALLVEPEGIRREETQNFLICMKDDVIS---QDVAIDLPTNVVEGSPRPSF 956

Qy        898 TAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFM 957
              : :||::|  :|   :   |::|:|: |||||:  ||||||:||||  |  :||::::|   |  : ::
Db        957 SVVGDIMGTAIQNVHQLLQMPFGNGEQNMVLFAPNIYVLDYLDKTRQLSEDVKSKTIGYL 1016

Qy        958 RQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWL 1017
                  ||::|  |:   |||:|:|  ||   |  |:||:|||: |  ||  |   :|  |  ||  |  :  ||
Db       1017 VSGYQKQLSYKHPDGSYSTFGIRDKEGNTWLTAFVYKSFAEASRFIYIDDNVQAQTLIWL 1076

Qy       1018 KGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLES 1077
                  ||::|  |:    |  ::::  ::||  ::  ::|:|||   :||          :  :: :  :  ||:
Db       1077 ATKQKTDGCFQSTGILVNNAMKGGVENELSLSAYITIALLEAGHSMSHTVIRNAFYCLET 1136

Qy       1078 EFSRGISDNYTLALITYALSSVGSPKAKEA-LNMLTWRAEQEGGMQFWVSSE------SK 1130
                : |:|  || ||:  ||    |    :    |:  |    |    |::   |  ::|   ::        |
```

```
Db       1137 ASEKNITDIYTQALVAYAFCLAGKAEICESFLRELQKSAKEVDGSKYWEQNQRSAPEKSH 1196

Qy       1131 LSDSWQPRSLDIEVAAYALLS-----HFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTT 1185
              | |  |  | |:|: :| ||:      :  |    ::   |::|: ||:|| ||||| |||
Db       1197 LLDHVQ--STDVEITSYVLLALLYKPNRSQEDLTKASAIVQWIIRQQNSYGGFASMQDTV 1254

Qy       1186 VALKALSEFAAL-----------MNTERTNIQV-TVTGPS----SPSPLAVVQPMAVNIS 1229
              |||:||: : |            :|:: |  :| ||   :       :|| | |   :::
Db       1255 VALQALAAYGAATYNSVTQNVIKINSKNTFEKVFTVNNENRLLLQQTPLPQV-PGKYSLT 1313

Qy       1230 ANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKEN-----KDDLNHVDLNV 1284
               || |  :  | ||:                     : ||    ::|: :    :|    |: :
Db       1314 VNGTGCVLIQTALRYNIHLP-----------EGAFGFSLSVQTSNASCPRDQPGKFDIVL 1362

Qy       1285 CTSFSGP-GRSGMALMEVNLLSGFMVPSEAISLSE-----TVKKVEYDHGKLNLYLDSV- 1337
               :|::|      | | :::| :|||| || :: || :         || :|||    : ||| ::
Db       1363 ISSYTGKRSSSNMVIIDVKMLSGF-VPVKS-SLDQLIDDHTVMQVEYKKNHVLLYLGNIL 1420

Qy       1338 ----NETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDV 1393
                   |  | |       ::| |:: : | | | ||||   || | |        || · |
Db       1421 QKRRKEVTFSVE----QDFVVTHPKPAPVQIYDYYETEEYAVAEYMS--------LCRGV 1468
```

```
RESULT 8
A2MG_MOUSE
ID   A2MG_MOUSE     STANDARD;      PRT;  1495 AA.
AC   Q61838; Q60628;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Alpha-2-macroglobulin precursor (Alpha-2-M).
GN   A2M.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 25-44 AND 1240-1259.
RC   TISSUE=Liver;
RX   MEDLINE=93076803; PubMed=1280217;
RA   van Leuven F., Torrekens S., Overbergh L., Lorent K., de Strooper B.,
RA   van den Berghe H.;
RT   "The primary sequence and the subunit structure of mouse alpha-2-
RT   macroglobulin, deduced from protein sequencing of the isolated
RT   subunits and from molecular cloning of the cDNA.";
RL   Eur. J. Biochem. 210:319-327(1992).
RN   [2]
RP   SEQUENCE OF 1-161 FROM N.A.
RC   STRAIN=129/J;
RX   MEDLINE=95095249; PubMed=7528166;
RA   Umans L., Serneels L., Hilliker C., Stas L., Overbergh L.,
RA   de Strooper B., van Leuven F., van den Berghe H.;
RT   "Molecular cloning of the mouse gene coding for alpha 2-macroglobulin
RT   and targeting of the gene in embryonic stem cells.";
RL   Genomics 22:519-529(1994).
CC   -!- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A
CC       UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,
CC       CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES
CC       FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT
CC       REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH
CC       TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST
CC       LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR
CC       WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE
CC       BAIT REGION A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE
CC       COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- DEVELOPMENTAL STAGE: CONTRARY TO THE RAT PROTEIN, WHICH IS AN
CC       ACUTE PHASE PROTEIN, THIS PROTEIN IS ALWAYS PRESENT AT HIGH LEVELS
CC       IN CIRCULATION.
CC   -!- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC       INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M93264; AAA39508.1; -.
DR   EMBL; U06977; AAA87890.1; -.
DR   PIR; S27001; S27001.
DR   HSSP; P01023; 1BV8.
DR   MGD; MGI:87854; A2m.
DR   GO; GO:0016975; F:alpha-2 macroglobulin; NAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW   Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
KW   Thioester bond.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25   1239       ALPHA-2-MACROGLOBULIN 165 kDa SUBUNIT.
FT   CHAIN      1240   1495       ALPHA-2-MACROGLOBULIN 35 kDa SUBUNIT.
FT   DOMAIN      686    744       BAIT REGION.
FT   DISULFID     48     86       BY SIMILARITY.
FT   DISULFID    249    298       BY SIMILARITY.
FT   DISULFID    267    286       BY SIMILARITY.
FT   DISULFID    277    430       BY SIMILARITY.
FT   DISULFID    594    783       BY SIMILARITY.
FT   DISULFID    642    689       BY SIMILARITY.
FT   DISULFID    833    861       BY SIMILARITY.
FT   DISULFID    859    895       BY SIMILARITY.
FT   DISULFID    933   1339       BY SIMILARITY.
FT   DISULFID   1092   1140       BY SIMILARITY.
FT   CROSSLNK    984    987       Isoglutamyl cysteine thioester (Cys-Gln)
FT                                (By similarity).
FT   CARBOHYD     55     55       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    157    157       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    382    382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    405    405       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    412    412       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    568    568       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    881    881       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    942    942       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1003   1003       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1385   1385       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1443   1443       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT     41     47       QESLKRP -> SGIPEKA (IN REF. 2).
FT   CONFLICT     91     91       I -> V (IN REF. 2).
FT   CONFLICT    142    144       VKF -> GII (IN REF. 2).
FT   CONFLICT    160    161       FP -> VS (IN REF. 2).
SQ   SEQUENCE   1495 AA;  165827 MW;  0D215688E80B4CCB CRC64;

  Query Match             18.7%;  Score 1373.5;  DB 1;  Length 1495;
  Best Local Similarity   27.0%;  Pred. No. 2.5e-67;
  Matches  418;  Conservative  271;  Mismatches  597;  Indels  263;  Gaps   48;

Qy         24 PGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQV-----------------TVKAELLKTA 66
              | | ||:    : |      |    : :    ||:|                   ||  |
Db          7 PTPAFLLLFLLLPRDATTATAKPQYVVLVPSEVYQESLKRPCVSLNHVNETVMLSLTLEY 66

Qy         67 SNLTVSVLEAEGVFEKGSF--KTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSF 124
              :     :|  : | :| ||        |:   ||            : : | ||     |    :
Db         67 AMQQTKLLTDQAV-DKDSFYCSPFTISGSPLPYT--FITVEIKGPTQ---RFIKKKSIQI 120

Qy        125 ETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQ 183
                 ||:|||| :|||  |  ||||:|::    |:|    :   :: |: || |  | || :
Db        121 IKAESPVFVQTDKPIYKPGQIVKFRVVSVDISFRPLNETFPVVYIETPKRNRIFQWQNIH 180

Qy        184 SDLGVISKTFQLSSHPILGDWSIQVQVND-QTYYQSFQVSEYVLPKFEVTLQTPLYCSMN 242
                 |:   :| ||   | ||  : :  || :  :          ||:| |||||||||  ::     :
Db        181 LAGGLHQLSFPLSVEPALGIYKVVVQKDSGKKIEHSFEVKEYVLPKFEVIIKMQKTMAFL 240
```

```
Qy    243 SKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFND----EE 298
            : |  |     ||||||| | |||            : : :      |  :  |      |
Db    241 EEELPITACGVYTYGKPVPGLVTL------------RVCRKYSRYRSTCHNQNSMSICAE 288

Qy    299 MKNVMDSSNGLSEYL-----DLSSPG---PVEILTTVTE-----SVTGISRNVSTNVFFK 345
                |     |: :        ||  |     :|:   :  |      :|||        |   |
Db    289 FSQQADDKGCFSQVVKTKVFQLSQKGHDMKIEVEAKIKEEGTGIELTGIGSCEIANALSK 348

Qy    346 QHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGS 405
                ::|     |   :|  | |:   | :     |  :          |:    |:    |    ::
Db    349 -----LKFTKVNTNYRPGLPFSGQVLLVDEKGKPIP----NKNITSVVSPLGYLSIFTTD 399

Qy    406 NSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHS--LFKSP 463
            |    : ::    |:|  |        :    :     : :     ::|     :  | ||   |   ||
Db    400 EHGLANI-SIDTSNFTAP-----FLRVVVTYKQNHVCYDNWWLDEFHTQADHSATLVFSP 453

Qy    464 SKTYIQLKT--------RDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNST 515
            |::||||:           : : |::         :  |::     |:: :||    ::|     |
Db    454 SQSYIQLELVFGTLACGQTQEIRIHYLLNEDIMKNEKTLTFYYLIKARG---SIGNLGSH 510

Qy    516 MFSLTPEN------------SWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLY 562
            : ||     |                |:|  :::|   |:  |  |:::|    :: |   ||: |
Db    511 VLSLEQGNMKGVFSLPIQVEPGMAPEAQLLIYAILPNEELVADAQNFEIEKCFANKVNLS 570

Qy    563 WSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLM----------------------- 599
            :    :: |:      |::        | |:  :  |||:|| |:
Db    571 FPSAQSLPASDTHLKVKAA-PLSLCALTAVDQSVLLLKPEAKLSPQSIYNLLPGKTVQGA 629

Qy    600 --------NASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTK--- 648
                     :   | |: |:: |     :|     :  ||     :: ::||   |: :  |:: :  |
Db    630 FFGVPVYKDHENCISGEDITHNGIVYTPKHSLG-DNDAHSIFQSVGINIFTNSKIHKPRF 688

Qy    649 -------DYIDGVYDNA---------------------EYAERFMEENEGHIVDIHDFS 679
                    :  ||    |                             :|::          :|:: :
Db    689 CQEFQHYPAMGGVAPQALAVAASGPGSSFRAMGVPMMGLDYSDEI-----NQVVEVRE-- 741

Qy    680 LGSSPHVRKHFPETWIW----LDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGL 735
                  |||:||||||         |   |  |||:||  |  |  |: :  |  |||
Db    742 -----TVRKYFPETWIWDLVPLDVSGD----GELAVKVPDTITEWKASAFCLSGTTGLGS 792

Qy    736 TTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMT 795
            ::|  :  ||||||||:  |  |||||:|||  |  |:  |:  ||:        :::|  :|  |   |
Db    793 SST-ISLQAFQPFFLELTLPYSVVRGEAFTLKATVLNYMSHCIQIRVDLEISPDF----- 846

Qy    796 SSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTA-------------- 841
               :      ||: :   :    :    || : :  |    |||:    | ||  :   :
Db    847 -LAVPVGGHENSHCICGNERKTVSWAVTPKSLGEVNFTRTAEALESQELCGNKLTEVPAL 905

Qy    842 --SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITA 899
                | |  :  ::|: ||||    :  :  ||      |    ||          |||  |  |||  |     :
Db    906 VHKDTVVKSVIVEPEGIEKEQTYNTLLCPQDTELQD---NSSLELPPNVVEGSARATHSV 962

Qy    900 IGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQ 959
            :||:||  ::     |   :|::||||||||||:   |    ||||:|:||   : :|||:   :|   ||::::
Db    963 LGDILGSAMQNLQNLLQMPYGCGEQNMVLFVPNIYVLNYLNETQQLTEAIKSKAINYLIS 1022

Qy    960 GYQRELLYQREDGSFSAFGNY---DPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTW 1016
            ||||:|  ||   |||:| |||:          :  |:|||:|||||:   :  :|  ||  :  :   |
Db   1023 GYQRQLNYQHSDGSYSTFGNHGGGNTPGNTWLTAFVLKAFAQAQSHIFIEKTHITNAFNW 1082

Qy   1017 LKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLE 1076
            |    ||  || |    |  |  :::: ::||     |||:|||   :||            :    |:  ::   ||
Db   1083 LSMKQKENGCFQQSGYLLNNAMKGGVDDEVTLSAYITIALLEMPLPVTHSAVRNALFCLE 1142

Qy   1077 SEFSRGISDN-----YTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFWVS---- 1126
            : ::  || :          || ||:  ||  :   |:   |   |   |     |   :|           |
Db   1143 TAWA-SISQSQESHVYTKALLAYAFALAGNKAKRSELLESLNKDAVKEEDSLHWQRPGDV 1201

Qy   1127 SESKLSDSWQPR--SLDIEVAAYALLSHFLQFQT-----------SEGIPIMRWLSRQRN 1173
            : |        :|||  |  ::|:  ||  ||::           :              |         |::|:|:|:|
Db   1202 QKVKALSFYQPRAPSAEVEMTAYVLLAYLTSESSRPTRDLSSSDLSTASKIVKWISKQQN 1261

Qy   1174 SLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPS-----------SPSPLAVVQ 1222
                 | ||    |||| |||:|||:: :      |  |:   :|  ||    |               | : | : :
Db   1262 SDGGLLLTQDTVVALQALSKYGSATFT-RSQKEVLVTSRSSGTFSKTFHVNSGNRLLLQE 1320

Qy   1223 PMAVNISAN------GFGFAICQLNVVYNVKASGSSRRRRSIQ-NQEAFDLDVAVKENKD 1275
                 : :   |           |    |   | :: ||:         :      ::|  |        :  |  |           :
Db   1321 VRLPDLPGNYVTKGSGSGCVYLQTSLKYNILPVADGKAPFALQVNTLPLNFDKA-----E 1375

Qy   1276 DLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGF--MVPS-EAISLSETVKKVEYDHGKLN 1331
                 |             : :      |::|        |  |  :::|  ::|||     |  ||  : :         :::  |   :      :
Db   1376 DHRTFQIRINVSYTGERPSSNMVIVDVKMVSGFIPMKPSVKRLQDQPNIQRTEVNTNHVL 1435

Qy   1332 LYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
             :|:: :           :         ::    |   |   :  |  :  :  ||||         |     |::
Db   1436 IYIEKLTNQTLGFSFAVEQDIPVKNLKPAPIKVYDYYETDEFTVEEYSA 1484
```

```
RESULT 9
CO3_LAMJA
ID   CO3_LAMJA     STANDARD;      PRT;  1673 AA.
AC   Q00685;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C3 precursor [Contains: C3A anaphylatoxin] (Fragment).
GN   C3.
OS   Lampetra japonica (Japanese lamprey) (Entosphenus japonicus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Lethenteron.
OX   NCBI_TaxID=94989;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=92251197; PubMed=1578150;
RA   Nonaka M., Takahashi M.;
RT   "Complete complementary DNA sequence of the third component of
RT   complement of lamprey. Implication for the evolution of thioester
RT   containing proteins.";
RL   J. Immunol. 148:3290-3295(1992).
CC   -!- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       COMPLEMENT SYSTEM. AFTER ACTIVATION (C3B), IT CAN BIND COVALENTLY,
CC       VIA ITS REACTIVE THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR
CC       IMMUNE AGGREGATES. CYCLOSTOMATES C3 APPEARS TO REPRESENT THE
CC       COMMON ANCESTOR OF MAMMALIAN C3 AND C4, SHOWING SIMILARITIES TO
CC       BOTH PROTEINS.
CC   -!- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D10087; BAA00983.1; -.
DR   PIR; I50806; I50806.
DR   HSSP; P01024; 1C3D.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; Macroglobln A2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Plasma; Inflammatory response; Glycoprotein;
```

```
KW   Signal; Thioester bond.
FT   NON_TER       1      1
FT   SIGNAL       <1     13       POTENTIAL.
FT   CHAIN        14   1673       COMPLEMENT C3.
FT   CHAIN        14    653       BETA CHAIN (BY SIMILARITY).
FT   CHAIN       657   1375       ALPHA CHAIN (BY SIMILARITY).
FT   CHAIN      1379   1673       GAMMA CHAIN (BY SIMILARITY).
FT   PEPTIDE     657    732       C3A ANAPHYLATOXIN (BY SIMILARITY).
FT   DOMAIN      678    713       ANAPHYLATOXIN-LIKE.
FT   DISULFID    678    705       BY SIMILARITY.
FT   DISULFID    679    712       BY SIMILARITY.
FT   DISULFID    692    713       BY SIMILARITY.
FT   CROSSLNK    986    989       Isoglutamyl cysteine thioester (Cys-Gln)
FT                                (By similarity).
SQ   SEQUENCE   1673 AA;  187767 MW;  4B5FC756B123A506 CRC64;

  Query Match             14.0%;  Score 1026.5;  DB 1;  Length 1673;
  Best Local Similarity   23.8%;  Pred. No. 3.1e-48;
  Matches  385;  Conservative  295;  Mismatches  619;  Indels  319;  Gaps   63;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPS----QVTVKAELLKTAS 67
Db          1 VLLLMSVGTSVTQDPMVLLSVPSVILIGSDVNV---LVDHAASTEDVSVVVRAEEFLTKK 57

Qy         68 NLTVSVLEAEGVFEKGSFKTLTLPSL-----------------PLNSADEIYELRVTGRT 110
Db         58 QLAT-------------QTITLTQLDPAIATLKLGFDIENPDKTNSASTKHHVRLVAKV 103

Qy        111 QDEILFSNSTR---LSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFK---PYKTSL 164
Db        104 ESKSFNKEITAHALLSYRSGH--VVVQTDKPIYTPDEKVKYRMFPMNREDVHRIPVRQSM 161

Qy        165 NILIKDPKSNLIQQWLS--QQSDLGVISKTF----QLSSHPILGDWSIQVQVNDQTYYQS 218
Db        162 TVDIVNADGVIVERQIKTIKATDEGIVDGTSFTIPAISKH---GTWKIFARMSGAPNINS 218

Qy        219 ---FQVSEYVLPKFEVTLQTPLYCSMNSK----HLNG-----TITAKYTYGKPVKGDVTL 266
Db        219 SAEFDVREYILPTFEV--------KINPKQRVFHINDEEFVVDITANYFNQELVSGTAYV 270

Qy        267 T-FLPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT 325
Db        271 RYFLENGDVPKLVDSSSTTLVAGEGLSILKKEKLLKLFPNAKDLLAF-------SLTIKT 323

Qy        326 TVTESVTGISRNVS-TNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEE 384
Db        324 TVLSSQAAETEEAELVGIKIVESRYQITATKTSRYFKPELPYFIQVEVRNADGS----PS 379

Qy        385 RRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLK 444
Db        380 KEVDVVAKVQVGSAT-----INPQKMRTDSNGLTSFTVTPPNVNQLTVTVRTDERH---- 430

Qy        445 AYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGN----KRLKELSYMV 500
Db        431 ---PSNEQGELVYTAQKYASASYMHIDV-TRIMRLGETLNVFLTAKTTQLNAVTHFTYMV 486

Qy        501 VSRGQLVAV--------GKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQ 552
Db        487 LTRGVIVKTNRKTKESGGGPSNVRIPITPD--MAPRFRFLAYYILPGGEIVADSVTVEVT 544

Qy        553 LVFKNKIKLYWSKVKAEPS--EKVSLRIS-VTQPDSIVGIVAVDKSVNLMNASNDITMEN 609
Db        545 ELCKSQVSL---SLKGRPTLEPKAMLTLDLIGEPDARVGLLAVDQAVYAVNRKHRLTQDR 601

Qy        610 VVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY--------- 660
Db        602 VWKAMETFDTGCTAEGGAGRPGVFSDAGLALITSKGL------NTTDRSEIGCPKVPSRK 655

Qy        661 ------------AERFMEENEGHIVD----------------------IHDF------ 678
Db        656 PRQLSMLQIRREAEKYTQEFRKCCVDGLKMSPTGQGCEERLKRVTGPKECVDAFLQCCKK 715

Qy        679 --------SLGSSPHVRK-----------------HFPETWIW--LDTNMGYRIYQEFE 710
Db        716 AEEYRKSESLGAKTVLRRNDFMELDLMNEDEVNMMAYFPQSWGWNKYKNSCKYGRHPQIR 775

Qy        711 VTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITI 770
Db        776 LQLPDTITTWNMQAVSISKTRGVCL-ADPLLLVSTKDFFIKLHLPYSVKRGEQTEIRVIL 834

Qy        771 FNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGE- 829
Db        835 YNYMEESL---TILTEMDIVESICSTSKSGAKPSQKS-TVKGKGAMVVSFPIVPLKIGEH 890

Qy        830 -IPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNT 888
Db        891 HISIRSRVYGRTFGDGVQKILRVAPEGVRDIRSESRSVHVEERETFFIKNEIS----PDV 946

Qy        889 VTGSERVQITAI-GDVLGPSI------NGLASLIRMPYGCGEQNMINFAPNIYILDYLTK 941
Db        947 VPNSDVLTFISVKGDELAETMVNCLDAKSISNLIQIPTGCGEQNMIKMAPTTLTLIYLDS 1006

Qy        942 KKQ---LTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLE 998
Db       1007 VQEWEKIGLHRREEAIGFLKQGYSRELSYRKADHSYAAF--IKRPSSTWLTAFVVKVYSL 1064

Qy        999 ADPYIDIDQNVLHRTYTW-LKGHQKSNGEFWDPGRVIHSELQ---GGNKSPVTLTAYIVT 1054
Db       1065 AKRVIIVDNQELCGPVEWIIKNRQNSDGSYREDGPVIHREMQGGVGGTEGHVSMTAFILI 1124

Qy       1055 SLLGYRKY----QPNI--DVQESIHFLESEFSRGISDNYTLALITYAL------SSVGSP 1102
Db       1125 GIQQAQEYCGVSVPNYKQSMNRAVQFLASKVS-DLKRMYTIAITRYALALQDPESEAAHS 1183

Qy       1103 KAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGI 1162
Db       1184 SWKKLENRTTFESK---GHRYWKAEET--SHVLRMSAISVEATAYGLLTYLRKKDYESAR 1238

Qy       1163 PIMRWLSRQRNSLGGFASTQDTTVALKALSEFAA-LMNTERTNIQVTVTGPSS------- 1214
Db       1239 EIVDWLTEQRNYGGGFQSTQDTILALQAMAQYKMDSSSKELIDVQLEITSPKNNFEKKMK 1298

Qy       1215 --------PSPLAVVQPMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDL 1266
Db       1299 ITEETRFVQEPHKIPPGGNITIKASGRGTFTLSIMSVFNKVAPSS-------KSCSTFDL 1351

Qy       1267 DVAVKENKD---------------------DLNHVD----LNVCTSFSG-----PGRSG 1295
Db       1352 KVTMTEADDGESPQGRLGWFDGKRRRRRDIGDEGGVEAVYRMNMCTRYKPRKEDLSSESG 1411

Qy       1296 MALMEVNLLSGFMVP--SEAISLSETVKK----VEYDHGKLNLYLDSVNETQ-FCVNIPA 1348
Db       1412 MTIIEVNMLTGF-IPDKNDLIQLKESVDKYISNYEITDSVLIIYWDKVPSTEDYCFAFKI 1470

Qy       1349 VRNFKVSNTQDASVSIVDYYEPRRQAVRSYN---SEVKLSS-C--DLCSDVQGCRPCE 1400
Db       1471 KQMLRSDMIQPVTASVYDYYSPADKCTRLYNLPGGYVELSPLCQNDLCQCVEVSCPAK 1528

RESULT 10
CO4_HUMAN
ID   CO4_HUMAN      STANDARD;      PRT;  1744 AA.
AC   P01028; Q13160; Q13906; Q14835; Q9NPK5; Q9UIP5;
DT   21-JUL-1986 (Rel. 01, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C4 precursor [Contains: C4A anaphylatoxin].
GN   C4A AND C4B.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (C4A AND C4B).
RC   TISSUE=Liver;
RX   MEDLINE=84156544; PubMed=6546707;
```

```
RA   Belt K.T., Carroll M.C., Porter R.R.;
RT   "The structural basis of the multiple forms of human complement
RT   component C4.";
RL   Cell 36:907-914(1984).
RN   [2]
RP   SEQUENCE FROM N.A. (C4A).
RX   MEDLINE=91108039; PubMed=1988494;
RA   Yu C.Y.;
RT   "The complete exon-intron structure of a human complement component
RT   C4A gene. DNA sequences, polymorphism, and linkage to the
RT   21-hydroxylase gene.";
RL   J. Immunol. 146:1057-1066(1991).
RN   [3]
RP   SEQUENCE FROM N.A. (C4B).
RC   TISSUE=Blood;
RX   MEDLINE=96163032; PubMed=8575831;
RA   Ulgiati D., Townend D.C., Christiansen F.T., Dawkins R.L.,
RA   Abraham L.J.;
RT   "Complete sequence of the complement C4 gene from the HLA-A1, B8,
RT   C4AQ0, C4B1, DR3 haplotype.";
RL   Immunogenetics 43:250-252(1996).
RN   [4]
RP   SEQUENCE FROM N.A. (C4B).
RA   Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
RA   Banta A., Swartzell S., Smith T.M., Spies T., Hood L.;
RT   "Sequence determination of 300 kilobases of the human class III MHC
RT   locus.";
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A. (C4A).
RA   Barlow K.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE OF 1-22 AND 1056-1225 FROM N.A.
RX   MEDLINE=85156269; PubMed=3838531;
RA   Belt K.T., Yu C.Y., Carroll M.C., Porter R.R.;
RT   "Polymorphism of human complement component C4.";
RL   Immunogenetics 21:173-180(1985).
RN   [7]
RP   SEQUENCE OF 680-756.
RX   MEDLINE=81264286; PubMed=6167582;
RA   Moon K.E., Gorski J.P., Hugli T.E.;
RT   "Complete primary structure of human C4a anaphylatoxin.";
RL   J. Biol. Chem. 256:8685-8692(1981).
RN   [8]
RP   SEQUENCE OF 957-1044.
RX   MEDLINE=82182029; PubMed=6978711;
RA   Campbell R.D., Gagnon J., Porter R.R.;
RT   "Amino acid sequence around the thiol and reactive acyl groups of
RT   human complement component C4.";
RL   Biochem. J. 199:359-370(1981).
RN   [9]
RP   SEQUENCE OF 990-1037.
RX   MEDLINE=82150875; PubMed=6950384;
RA   Harrison R.A., Thomas M.L., Tack B.F.;
RT   "Sequence determination of the thiolester site of the fourth
RT   component of human complement.";
RL   Proc. Natl. Acad. Sci. U.S.A. 78:7388-7392(1981).
RN   [10]
RP   SEQUENCE OF 1-21 FROM N.A.
RX   MEDLINE=94282044; PubMed=8012361;
RA   Sargent C.A., Anderson M.J., Hsieh S.L., Kendall E.,
RA   Gomez-Escobar N., Campbell R.D.;
RT   "Characterisation of the novel gene G11 lying adjacent to the
RT   complement C4A gene in the human major histocompatibility complex.";
RL   Hum. Mol. Genet. 3:481-488(1994).
RN   [11]
RP   SEQUENCE OF 1405-1431, AND SULFATION.
RX   MEDLINE=86111851; PubMed=3944109;
RA   Hortin G., Sims H., Strauss A.W.;
RT   "Identification of the site of sulfation of the fourth component of
RT   human complement.";
RL   J. Biol. Chem. 261:1786-1793(1986).
RN   [12]
RP   STRUCTURAL BASIS OF POLYMORPHISM.
RX   MEDLINE=87080272; PubMed=2431902;
RA   Yu C.Y., Belt K.T., Giles C.M., Campbell R.D., Porter R.R.;
RT   "Structural basis of the polymorphism of human complement components
RT   C4A and C4B: gene size, reactivity and antigenicity.";
RL   EMBO J. 5:2873-2881(1986).
RN   [13]
RP   VARIANT C4A6 ALLOTYPE.
RX   MEDLINE=92242905; PubMed=1573268;
RA   Anderson M.J., Milner C.M., Cotton G.H., Campbell R.D.;
RT   "The coding sequence of the hemolytically inactive C4A6 allotype of
RT   human complement component C4 reveals that a single arginine to
RT   tryptophan substitution at beta-chain residue 458 is the likely cause
RT   of the defect.";
RL   J. Immunol. 148:2795-2802(1992).
CC   -!- FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY
CC       ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A
CC       ANAPHYLATOXIN.
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C4,
CC       C4A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES.
CC   -!- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR
CC       AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER
CC       OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).
CC   -!- POLYMORPHISM: HUMAN COMPLEMENT COMPONENT C4 IS POLYMORPHIC WITH AT
CC       LEAST TWO LOCI, C4A AND C4B. 13 ALLELES OF C4A AND 22 ALLELES OF
CC       C4B HAVE BEEN DETECTED. THE ALLELE SHOWN HERE IS C4A4.
CC   -!- POLYMORPHISM: THE C4A ALLELES CARRY THE BLOOD GROUP RODGERS WHILE
CC       THE C4B ALLELES CARRY THE BLOOD GROUP CHIDO.
CC   -!- DISEASE: Defects in C4A are the cause of C4A deficiency
CC       [MIM:120810].
CC   -!- DISEASE: The C4A6 allotype is totally deficient in hemolytic
CC       activity.
CC   -!- MISCELLANEOUS: C4A ALLOTYPES REACT MORE RAPIDLY WITH THE AMINO
CC       GROUP OF PEPTIDE ANTIGENS WHILE C4B ALLOTYPES REACT MORE RAPIDLY
CC       WITH THE HYDROXYL GROUP OF CARBOHYDRATE ANTIGENS.
CC   -!- MISCELLANEOUS: C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III
CC       PROTEIN.
CC   -!- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; K02403; AAB59537.1; -.
DR   EMBL; K02404; AAA59651.1; -.
DR   EMBL; M59815; AAA51855.1; -.
DR   EMBL; M59816; AAA51855.1; JOINED.
DR   EMBL; U24578; AAA99717.1; -.
DR   EMBL; AF019413; AAB67980.1; -.
DR   EMBL; AL049547; CAB89302.1; -.
DR   EMBL; M14823; AAA35617.1; -.
DR   EMBL; M14824; AAA52292.1; -.
DR   EMBL; X77491; CAA54627.1; -.
DR   PIR; I56095; C4HU.
DR   HSSP; P01031; 1KJS.
DR   SWISS-2DPAGE; P01028; HUMAN.
DR   Genew; HGNC:1323; C4A.
DR   Genew; HGNC:1324; C4B.
DR   MIM; 120810; -.
DR   MIM; 120820; -.
DR   MIM; 120790; -.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0006958; P:complement activation, classical pathway; NAS.
```

```
DR   GO; GO:0006954; P:inflammatory response; NAS.
DR   GO; GO:0006937; P:regulation of muscle contraction; NAS.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; MacroglobInA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Plasma; Glycoprotein; Sulfation; Signal;
KW   Inflammatory response; Polymorphism; Disease mutation;
KW   Blood group antigen; Thioester bond.
FT   SIGNAL        1     19
FT   CHAIN        20    675       COMPLEMENT C4, BETA CHAIN.
FT   PROPEP      676    679
FT   CHAIN       680   1446       COMPLEMENT C4, ALPHA CHAIN.
FT   PROPEP     1447   1453
FT   CHAIN      1454   1744       COMPLEMENT C4, GAMMA CHAIN.
FT   PEPTIDE     680    756       C4A ANAPHYLATOXIN.
FT   DOMAIN      702    736       ANAPHYLATOXIN-LIKE.
FT   DISULFID    702    728       BY SIMILARITY.
FT   DISULFID    703    735       BY SIMILARITY.
FT   DISULFID    716    736       BY SIMILARITY.
FT   CROSSLNK   1010   1013       Isoglutamyl cysteine thioester (Cys-Gln).
FT   MOD_RES    1417   1417       SULFATION.
FT   MOD_RES    1420   1420       SULFATION.
FT   MOD_RES    1422   1422       SULFATION.
FT   CARBOHYD    226    226       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    862    862       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1328   1328       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   1391   1391       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     477    477       R -> W (in allotype C4A6).

  Query Match             13.5%;  Score 991.5;  DB 1;  Length 1744;
  Best Local Similarity   24.7%;  Pred. No. 2.8e-46;
  Matches  409;  Conservative  250;  Mismatches  636;  Indels  359;  Gaps   65;

Qy         26 PRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTAS--NLTVSVLEAEGVFEKG 83
Db         21 PRLLLFSPSVVHLGVPLSVGVQLQDVPRGQVVKGSVFLRNPSRNNVPCSPKVDFTLSSER 80

Qy         84 SFKTLTLPSLPLNSA-----------DEIYELRVTGRTQDEILFSNST-----RLSFETK 127
Db         81 DFALLSL-QVPLKDAKSCGLHQLLRGPEVQLVAHSPWLKDSL--SRTTNIQGINLLFSSR 137

Qy        128 RISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLG 187
Db        138 RGHLFLQTDQPIYNPGQRVRYRVFALDQKMRPSTDTITVMVENSHGLRV-----RKKEVY 192

Qy        188 VISKTFQLS------SHPILGDWSIQVQVNDQTYYQS---FQVSEYVLPKFEVTLQTP-- 236
Db        193 MPSSIFQDDFVIPDISEP--GTWKISARFSDGLESNSSTQFEVKKYVLPNFEVKI-TPGK 249

Qy        237 LYCSMNSKHLNG---TITAKYTYGKPVKGDVTLTFLPLSFWGKKKNI----TKTFKINGS 289
Db        250 PYILTVPGHLDEMQLDIQARYIYGKPVQGVAYVRFGLLDEDGKKTFFRGLESQTKLVNGQ 309

Qy        290 ANFSFNDEEMKNVMDSSN-GLSEYLDLSSPGPVEILTTVTESVTGISRNVS-TNVFFKQH 347
Db        310 SHISLSKAEFQDALEKLNMGITDLQGLR----LYVAAAIIESPGGEMEEAELTSWYFVSS 365

Qy        348 DYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNS 407
Db        366 PFSLDLSKTKRHLVPGAPFLLQALVREMSGSP------ASGIPVKV---------SATVS 410

Qy        408 GNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPS--K 465
Db        411 SPGSVPEVQDIQQNTDGSGQVSIPIIIPQTISELQLSVSAGSPHPAIARLTVAAPPSGGP 470

Qy        466 TYIQLKTRDEN-IKVGSPFEL---VVSGNKRLKELSYMVVSRGQLVAVG---KQNSTMFS 518
Db        471 GFLSIERPDSRPPRVGDTLNLNLRAVGSGATFSHYYYMILSRGQIVFMNREPKRTLTSVS 530

Qy        519 LTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQL-VFKNKIKLYWSKVKA-EPSEKVSL 576
Db        531 VFVDHHLAPSFYFVAFYYHGD-HPVANSLRVDVQAGACEGKLELSVDGAKQYRNGESVKL 589

Qy        577 RISVTQPDSIVGIVAVDKSVNLMNASN--DITMENVVHELELYNTGYYLGMFMNSFAVFQ 634
Db        590 HLE-TDSLALVALGALDTALYAAGSKSHKPLNMGKVFEAMNSYDLGCGPGGGDSALQVFQ 648

Qy        635 ECGL-------WVLT---------------DANLTK-------DYI---------DGV- 654
Db        649 AAGLAFSDGDQWTLSRKRLSCPKEKTTRKKRNVNFQKAINEKLGQYASPTAKRCCQDGVT 708

Qy        655 -----------------------YDNAEYAERFMEEN----------------EGHIVD 674
Db        709 RLPMMRSCEQRAARVQQPDCREPFLSCCQFAESLRKKSRDKGQAGLQRALEILQEEDLID 768

Qy        675 IHDFSLGSSPHVRKHFPETWIWLDTNMGYRI-----YQEFEVTVPDSITSWVATGFVISE 729
Db        769 EDDIP------VRSFFPENWLW-------RVETVDRFQILTLWLPDSLTTWEIHGLSLSK 815

Qy        730 DLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDK 789
Db        816 TKGL-CVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKNLTVSVHVSPVEG 874

Qy        790 FDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSP---TASDAVT 846
Db        875 LCL------AGGGLAQQVLVPAGSARPVAFSVVPTAAAAVSLKVVARGSFEFPVGDAVS 928

Qy        847 QMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSF--SFPPNTVTGSE---RVQITAIG 901
Db        929 KVLQIEKEG--AIHREELVYEL--NPLDHRGRTLEIPGNSDPNMIPDGDFNSYVRVTASD 984

Qy        902 --DVLGP----SINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQ---LTDNLKEK 952
Db        985 PLDTLGSEGALSPGGVASLLRLPRGCGEQTMIYLAPTLAASRYLDKTEQWSTLPPETKDH 1044

Qy        953 ALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHR 1012
Db       1045 AVDLIQKGYMRIQQFRKADGSYAAWLSRD--SSTWLTAFVLKVLSLAQEQVGGSPEKLQE 1102

Qy       1013 TYTWLKGHQKSNGEFWDPGRVIHSELQG---GNKSPVTLTAYIVTSL-LGYRKYQP---- 1064
Db       1103 TSNWLLSQQQADGSFQDPCPVLDRSMQGGLVGNDETVALTAFVTIALHHGLAVFQDEGAE 1162

Qy       1065 ------NIDVQESIHFLESEFSRGISDNYTLALITYALSSVGSP--KAKEALNMLTWRAE 1116
Db       1163 PLKQRVEASISKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQ 1222

Qy       1117 QEGGMQFWVSSESKLSDSWQPR--------------SLDIEVAAYALLSHFLQFQTSEGI 1162
Db       1223 ETGDNLYWGSVTGSQSNAVSPTPAPRNPSDPMPQAPALWIETTAYALL-HLL---LHEGK 1278

Qy       1163 PIM-----RWLSRQRNSLGGFASTQDTTVALKALSEF-AALMNTERTNIQVTVTGPSSPS 1216
Db       1279 AEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERGLNVTLS------ 1332

Qy       1217 PLAVVQPMAVNISANGFGFAICQLN----------------VVYNVKASGSSRRRRSI-- 1258
Db       1333 ----------STGRNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLR 1382

Qy       1259 -------QNQEAFDLDVAV------------KENKDDLNHVDL----------------- 1282
Db       1383 TYNVLDMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQ 1442

Qy       1283 -------------------------NVCTSFSGP-GRSGMALMEVNLLSGFMV----PS 1311
```

```
Db       1443 LFEGRRNRRRREAPKVVEEQESRVHYTVCIWRNGKVGLSGMAIADVTLLSGFHALRADLE 1502

Qy       1312 EAISLSET-VKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEP 1370
                :  |||:  |     | :    : || |||  :: ||    ||:   |     | || :: ||| |
Db       1503 KLTSLSDRYVSHFETEGPHVLLYFDSVPTSRECVGFEAVQEVPVGLVQPASATLYDYYNP 1562

Qy       1371 RRQAVRSYNSEVK------LSSCDLCSDVQGCRP 1398
                 |:    | :  |         | | ::|    :| |
Db       1563 ERRCSVFYGAPSKSRLLATLCSAEVCQCAEGKCP 1596


RESULT 11
CO4_MOUSE
ID   CO4_MOUSE      STANDARD;      PRT;  1738 AA.
AC   P01029; Q61859;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C4 precursor [Contains: C4A anaphylatoxin].
GN   C4.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85298264; PubMed=3862104;
RA   Sepich D.S., Noonan D.J., Ogata R.T.;
RT   "Complete cDNA sequence of the fourth component of murine
RT   complement.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:5895-5899(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B12.WR;
RX   MEDLINE=87309760; PubMed=3624865;
RA   Rosa P.A., Sepich D.S., Robins D.M., Ogata R.T.;
RT   "Constitutive expression of Slp genes in mouse strain B10.WR directed
RT   by C4 regulatory sequences.";
RL   J. Immunol. 139:1568-1577(1987).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B12.WR; TISSUE=Liver;
RX   MEDLINE=89380278; PubMed=2777798;
RA   Ogata R.T., Rosa P.A., Zepf N.E.;
RT   "Sequence of the gene for murine complement component C4.";
RL   J. Biol. Chem. 264:16565-16572(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FM; TISSUE=Liver;
RX   MEDLINE=85289294; PubMed=2993295;
RA   Nonaka M., Nakayama K., Yeul Y.D., Takahashi M.;
RT   "Complete nucleotide and derived amino acid sequences of the fourth
RT   component of mouse complement (C4). Evolutionary aspects.";
RL   J. Biol. Chem. 260:10936-10943(1985).
RN   [5]
RP   SEQUENCE OF 651-810 AND 924-1083 FROM N.A.
RX   MEDLINE=85038607; PubMed=6208559;
RA   Nonaka M., Takahashi M., Natsuume-Sakai S., Nonaka M., Tanaka S.,
RA   Shimizu A., Honjo T.;
RT   "Isolation of cDNA clones specifying the fourth component of mouse
RT   complement and its isotype, sex-limited protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:6822-6826(1984).
RN   [6]
RP   SEQUENCE OF 1099-1142 FROM N.A.
RC   STRAIN=B10.WR, C57BL/6, C3H/He, CBA/J, B10.BR, and DBA2;
RX   MEDLINE=90353398; PubMed=2387317;
RA   Ogata R.T., Zepf N.E.;
RT   "C4 from C4-high and C4-low mouse strains have identical sequences in
RT   the region corresponding to the isotype-specific segment of human
RT   C4.";
RL   Eur. J. Immunol. 20:1607-1610(1990).
RN   [7]
RP   SEQUENCE OF 1105-1449 FROM N.A.
RX   MEDLINE=85166208; PubMed=3856857;
RA   Levi-Strauss M., Tosi M., Steinmetz M., Klein J., Meo T.;
RT   "Multiple duplications of complement C4 gene correlate with H-2-
RT   controlled testosterone-independent expression of its sex-limited
RT   isoform, C4-Slp.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:1746-1750(1985).
RN   [8]
RP   SEQUENCE OF 1257-1376 FROM N.A.
RX   MEDLINE=85038859; PubMed=6149581;
RA   Tosi M., Levi-Strauss M., Duponchel C., Meo T.;
RT   "Sequence heterogeneity of murine complementary DNA clones related to
RT   the C4 and C4-Slp isoforms of the fourth complement component.";
RL   Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:389-394(1984).
RN   [9]
RP   SEQUENCE OF 1360-1511 FROM N.A.
RX   MEDLINE=83273751; PubMed=6192448;
RA   Ogata R.T., Shreffler D.C., Sepich D.S., Lilly S.P.;
RT   "cDNA clone spanning the alpha-gamma subunit junction in the
RT   precursor of the murine fourth complement component (C4).";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:5061-5065(1983).
RN   [10]
RP   SEQUENCE OF 1-128 FROM N.A.
RC   STRAIN=FM; TISSUE=Liver;
RX   MEDLINE=86031969; PubMed=2997024;
RA   Nonaka M., Nakayama K., Yeul Y.D., Shimizu A., Takahashi M.;
RT   "Molecular cloning and characterization of complementary and genomic
RT   DNA clones for mouse C4 and Slp.";
RL   Immunol. Rev. 87:81-99(1985).
RN   [11]
RP   SEQUENCE OF 1-21 FROM N.A.
RX   MEDLINE=87017050; PubMed=3464002;
RA   Nonaka M., Kimura H., Yeul Y.D., Yokoyama S., Nakayama K.,
RA   Takahashi M.;
RT   "Identification of the 5'-flanking regulatory region responsible for
RT   the difference in transcriptional control between mouse complement C4
RT   and Slp genes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:7883-7887(1986).
CC   -!- FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY
CC       ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A
CC       ANAPHYLATOXIN.
CC   -!- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR
CC       AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER
CC       OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).
CC   -!- MISCELLANEOUS: C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III
CC       PROTEIN.
CC   -!- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; K00019; AAA39554.1; -.
DR   EMBL; M11729; AAA39506.1; -.
DR   EMBL; M12968; AAA39558.1; -.
DR   EMBL; M12970; AAA39555.1; -.
DR   EMBL; M12972; AAA39556.1; -.
DR   EMBL; M12969; AAA39559.1; -.
DR   EMBL; M11789; AAA39557.1; -.
DR   EMBL; K02798; AAC42021.1; -.
DR   EMBL; M17440; AAA39561.1; -.
DR   EMBL; M14225; AAA39563.1; -.
DR   EMBL; M14226; AAA39684.1; -.
DR   EMBL; X55493; CAA39112.1; -.
DR   EMBL; X55495; CAA39114.1; -.
DR   PIR; A24558; A24558.
DR   PIR; A29176; A29176.
DR   HSSP; P01031; 1KJS.
```

```
DR   MGD; MGI:88228; C4.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; Macroglobln A2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Plasma; Glycoprotein; Sulfation; Signal;
KW   Inflammatory response; Thioester bond.
FT   SIGNAL        1     19
FT   CHAIN        20    673       COMPLEMENT C4, BETA CHAIN.
FT   PROPEP      674    677
FT   CHAIN       678   1443       COMPLEMENT C4, ALPHA CHAIN.
FT   PROPEP     1444   1447
FT   CHAIN      1448   1738       COMPLEMENT C4, GAMMA CHAIN.
FT   PEPTIDE     678    753       C4A ANAPHYLATOXIN.
FT   DOMAIN      700    734       ANAPHYLATOXIN-LIKE.
FT   DISULFID    700    726       BY SIMILARITY.
FT   DISULFID    701    733       BY SIMILARITY.
FT   DISULFID    714    734       BY SIMILARITY.
FT   CROSSLNK   1006   1009       Isoglutamyl cysteine thioester (Cys-Gln)
FT                                (By similarity).
FT   MOD_RES    1413   1413       SULFATION.
FT   MOD_RES    1416   1416       SULFATION.
FT   MOD_RES    1417   1417       SULFATION.
FT   CARBOHYD    224    224       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    743    743       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   1387   1387       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    132    132       F -> Y (IN REF. 4).
FT   CONFLICT    327    327       G -> E (IN REF. 4).
FT   CONFLICT    570    570       Q -> E (IN REF. 4).
FT   CONFLICT    720    720       R -> G (IN REF. 5).
FT   CONFLICT    739    740       DL -> AI (IN REF. 5).
FT   CONFLICT    838    838       P -> R (IN REF. 4).
FT   CONFLICT    993    993       P -> L (IN REF. 5).
FT   CONFLICT   1043   1043       D -> E (IN REF. 5).
FT   CONFLICT   1119   1119       V -> A (IN REF. 7).
FT   CONFLICT   1190   1190       A -> T (IN REF. 7).
FT   CONFLICT   1324   1324       K -> N (IN REF. 4).
FT   CONFLICT   1401   1401       G -> S (IN REF. 9).
FT   CONFLICT   1442   1442       R -> K (IN REF. 4).
FT   CONFLICT   1453   1453       A -> V (IN REF. 4).
SQ   SEQUENCE   1738 AA;  192870 MW;  D1EE02AE7AB42BFF CRC64;

  Query Match            13.4%;  Score 987;  DB 1;  Length 1738;
  Best Local Similarity  24.9%;  Pred. No. 4.8e-46;
  Matches  415;  Conservative  245;  Mismatches  643;  Indels  366;  Gaps   64;

Qy         10 AHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAEL------- 62
Db          8 AWVFSFCASSL---QKPRLLLFSPSVVNLGTPLSVGVQLLDAPPGQ-EVKGSVFLRNPKG 63
Qy         63 ----------LKTASNLTVSVLEA--EGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRT 110
Db         64 GSCSPKKDFKLSSGDDFVLLSLEVPLEDVRSCGLFDLRRAPHIQLVAQSP--WLRNTAFK 121
Qy        111 QDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKD 170
Db        122 ATE---TQGVNLLFSSRRGHIFVQTDQPIYNPGQRVRYRVFALDQKMRPSTDFLTITVEN 178
Qy        171 PKS-NLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVND---QTYYQSFQVSEYVL 226
Db        179 SHGLRVLKKEIFTSTSIFQDAFTIPDISEP--GTWKISARFSDGLESNRSTHFEVKKYVL 236
Qy        227 PKFEVTLQTP-----LYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNI- 280
Db        237 PNFEVKI-TPWKPYILMVPSNSDEIQLDIQARYIYGKPVQGVAYTRFALMDEQGKRTFLR 295
Qy        281 ---TKTFKINGSANFSFNDEEMKNVMDSSN-------GLSEYLDLSSPGPVEILTTVTES 330
Db        296 GLETQAKLVEGRTHISISKDQFQAALDKINIGVRDLEGLRLY---------AATAVIES 345
Qy        331 VTGISRNVS-TNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNV 389
Db        346 PGGEMEEAELTSWRFVSSAFSLDLSRTKRHLVPGAHFLLQALVQEMSGSE------ASNV 399
Qy        390 VITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLG 449
Db        400 PVKVS----ATLVSGSDS--QVLDIQQSTN----GIGQVSISFPIPPTVTELRLLVSAGS 449
Qy        450 SKSSMAVHSLFKSPSK--TYIQLKTRD-ENIKVGSPFEL----VVSGNKRLKELSYMVVS 502
Db        450 LYPAIARLTVQAPPSRGTGFLSIEPLDPRSPSVGDTFILNLQPVGIPAPTFSHYYYMIIS 509
Qy        503 RGQLVAVGKQ---NSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQ---LVFK 556
Db        510 RGQIMAMGREPRKTVTSVSVLVDHQLAPSFYFVAYFYH-QGHPVANSLLINIQSRDCEGK 568
Qy        557 NKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLM--NASNDITMENVVHEL 614
Db        569 LQLKVDGAK-EYRNADMMKLRIQ-TDSKALVALGAVDTALYAVGGRSHKPLDMSKVFEVI 626
Qy        615 ELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYID--------------------- 652
Db        627 NSYNVGCGPGGGDDALQVFQDAGLAFSDGDRLTQTREDLSCPKEKKSRQKRNVNFQKAVS 686
Qy        653 ---GVYDN------------------------------------------AEYAERFMEE 667
Db        687 EKLGQYSSPDAKRCCQDGMTKLPMKRTCEQRAARVPQQACREPFLSCCKFAEDLRRNQTR 746
Qy        668 NEGH-------------IVDIHDFSLGSSPHVRKHFPETWIW----LDTNMGYRIYQEFE 710
Db        747 SQAHLARNNHNMLQEEDLIDEDDIL------VRTSFPENWLWRVEPVDSS------KLLT 794
Qy        711 VTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITI 770
Db        795 VWLPDSMTTWEIHGVSLSKSKGL-CVAKPTRVRVFRKFHLHLRLPISIRRFEQFELRPVL 853
Qy        771 FNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEI 830
Db        854 YNYLND--DVAVSVHVTPVEGLCLAGGGMMA----QQVTVPAGSARPVAFSVVPTAAANV 907
Qy        831 PITVTALSP-TASDAVTQMILVKAEGIEKSYSQSILLDLTD-NRLQSTLKTLSFSFPPN- 887
Db        908 PLKVVARGVFDLGDAVSKILQIEKEG--AIHREELVYNLDPLNNLGRTLE-IPGSSDPNI 964
Qy        888 ----------TVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILD 937
Db        965 VPDGDFSSLVRVTASEPLETMGSEGALSP--GGVASLLRLPQGCAEQTMIYLAPTLTASN 1022
Qy        938 YLTKKKQ---LTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLR 994
Db       1023 YLDRTEQWSKLSPETKDHAVDLIQKGYMRIQQFRKNDGSFGAWLHRD--SSTWLTAFVLK 1080
Qy        995 CFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQG---GNKSPVTLTAY 1051
Db       1081 ILSLAQEQVGNSPEKLQETASWLLAQQLGDGSFHDPCPVIHRAMQGGLVGSDETVALTAF 1140
Qy       1052 IVTSL-----------LGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVG 1100
Db       1141 VVIALHHGLDVFQDDDAKQLKNRVEASITKANSFLGQKASAGLLGAHAAAITAYALTLT- 1199
Qy       1101 SPKAKEAL-----NMLTWRAEQEGGMQFW---VSSESK-LSDSWQPRS----------LD 1141
Db       1200 --KASEDLRNVAHNSLMAMAEETGEHLYWGLVLGSQDKVVLRPTAPRSPTEPVPQAPALW 1257
Qy       1142 IEVAAYALLSHFLQFQTSEGIPIM-----RWLSRQRNSLGGFASTQDTTVALKALSEF-A 1195
```

```
Db       1258 IETTAYALLHLLLR----EGKGKMADKAASWLTHQGSFHGAFRSTQDTVVTLDALSAYWI 1313

Qy       1196 ALMNTERTNIQVT-------------------VTGPSSPSPLAVVQPMAVNISANGFGF 1235
              |  ||   ::||                    | |     ::   ::| :  |  |
Db       1314 ASHTTEEKALKVTLSSMGRNGLKTHGLHLNNHQVKGLEEELKFSLGSTISVKVEGNSKGT 1373

Qy       1236 -------------AICQLNVVYNVKASGS----------------------------- 1251
                           || ::   || :|:
Db       1374 LKILRTYNVLDMKNTTCQ-DLQIEVKVTGAVEYAWDANEDYEDYYDMPAADDPSVPLQPV 1432

Qy       1252 ----------SRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGP-GRSGMALME 1300
                        |||||          | |   ::  :  |    ||    :|  | ||||: :
Db       1433 TPLQLFEGRRSRRRRE-----------APKVAEEQESRVQYTVCIWRNGKLGLSGMAIAD 1481

Qy       1301 VNLLSGFMV----PSEAISLSET-VKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVS 1355
              : |||||      :  |||:  |   |  |   : || |||  |: ||    | :   |
Db       1482 ITLLSGFHALRADLEKLTSLSDRYVSHFETDGPHVLLYFDSVPTTRECVGFGASQEVVVG 1541

Qy       1356 NTQDASVSIVDYYEPRRQAVRSYNSEVK------LSSCDLCSDVQGCRP 1398
                |  :|  :  ||| |   :      |  :  |      | | |:|    :|  |
Db       1542 LVQPSSAVLYDYYSPDHKCSVFYAAPTKSQLLATLCSGDVCQCAEGKCP 1590
```

```
RESULT 12
CO3_ONCMY
ID   CO3_ONCMY      STANDARD;      PRT;  1640 AA.
AC   P98093;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C3-1 [Contains: C3A anaphylatoxin] (Fragment).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Liver;
RX   MEDLINE=94065166; PubMed=8245455;
RA   Lambris J.D., Lao Z., Pang J., Alsenz J.;
RT   "Third component of trout complement. cDNA cloning and conservation
RT   of functional sites.";
RL   J. Immunol. 151:6123-6134(1993).
CC   -!- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
CC       REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.
CC       AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE
CC       THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.
CC   -!- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG
CC       RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE
CC       BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,
CC       RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA'
CC       CHAIN).
CC   -!- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L24433; AAB05029.1; ALT_INIT.
DR   PIR; I51339; I51339.
DR   HSSP; P01024; 1C3D.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; MacrogloblnA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Complement alternate pathway; Plasma;
KW   Inflammatory response; Glycoprotein; Thioester bond.
FT   NON_TER       1      1
FT   CHAIN         1   1640       COMPLEMENT C3.
FT   CHAIN         1    642       BETA CHAIN.
FT   CHAIN       647   1640       ALPHA CHAIN.
FT   PEPTIDE     647    722       C3A ANAPHYLATOXIN.
FT   CHAIN       723   1640       C3B (ALPHA' CHAIN).
FT   PEPTIDE     723    931       C3C FRAGMENT (BY SIMILARITY).
FT   PEPTIDE     932   1278       C3DG FRAGMENT (BY SIMILARITY).
FT   PEPTIDE     932   1033       C3G FRAGMENT (BY SIMILARITY).
FT   PEPTIDE    1034   1278       C3D FRAGMENT (BY SIMILARITY).
FT   PEPTIDE    1279   1295       C3F FRAGMENT (BY SIMILARITY).
FT   SITE        722    723       CLEAVAGE (BY C3 CONVERTASE).
FT   SITE        931    932       CLEAVAGE (BY FACTOR I) (BY SIMILARITY).
FT   SITE       1278   1279       CLEAVAGE (BY FACTOR I) (BY SIMILARITY).
FT   SITE       1295   1296       CLEAVAGE (BY FACTOR I).
FT   DOMAIN      668    703       ANAPHYLATOXIN-LIKE.
FT   DOMAIN     1420   1430       PROPERDIN-BINDING.
FT   DISULFID    536    797       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    603    638       BY SIMILARITY.
FT   DISULFID    668    695       BY SIMILARITY.
FT   DISULFID    669    702       BY SIMILARITY.
FT   DISULFID    682    703       BY SIMILARITY.
FT   DISULFID    853   1488       BY SIMILARITY.
FT   DISULFID   1079   1135       BY SIMILARITY.
FT   DISULFID   1335   1464       BY SIMILARITY.
FT   DISULFID   1481   1486       BY SIMILARITY.
FT   DISULFID   1493   1563       BY SIMILARITY.
FT   DISULFID   1510   1638       BY SIMILARITY.
FT   DISULFID   1614   1623       BY SIMILARITY.
FT   CARBOHYD    164    164       N-LINKED (GLCNAC. . .).
FT   CROSSLNK    988    991       Isoglutamyl cysteine thioester (Cys-Gln)
FT                                (By similarity).
SQ   SEQUENCE   1640 AA;  182104 MW;  0965B4FAF1E87812 CRC64;

  Query Match             13.4%;  Score 984.5;  DB 1;  Length 1640;
  Best Local Similarity   24.1%;  Pred. No. 6.1e-46;
  Matches  387;  Conservative  298;  Mismatches  571;  Indels  349;  Gaps   67;

Qy         18 AALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEAE 77
              |||  |         ::||  ::|  |  |   |  ||   :|      :  ||   ::|     :    :
Db          1 AALQV-------LSAPNLLRVGSNENIFVESQDHVGGPLNVKI-MVKNHPTQSKELASKS 52

Qy         78 GVFEK-GSFKTLT--------LPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKR 128
               | ::  :|: :|         |        |           : : :    |   ::::     :||::
Db         53 VVLDQANNFQAMTQLVIQRGPLVDDPKQKQYVVLQAQFPDRLLEKVVL-----VSFQSGY 107

Qy        129 ISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY-------------KTSLNILIKDPKSN 174
               |  |||||| :| |    | :|: ::    :|                 : ::::  |   |::
Db        108 I--FIQTDKTIYTPASTVHYRVFSMTPGLEPLTREIFEDQEVAKNKEIAVSVEIMTPENI 165

Qy        175 LIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVN---DQTYYQSFQVSEYVLPKFEV 231
               | : :     | || |   |:|         | |  :   :         :|:    |:| ||||| |||
Db        166 TIFREI-VNPDKGVKSGQFKLPDIVSFGTWHVVTRFQSTPQKTFSSEFEVKEYVLPSFEV 224

Qy        232 TLQTP--LYCSMNSKHLNGTITAKYTYGKPVKGDVTLTF-----------LPLSFWG--- 275
              :| ||    :   ::      |  |||:| ||| | |       :  |                  | |
Db        225 SL-TPAKAFFYVDDNDLTVDITARYLYGKEVTGTGYVVFGVITTESEKKSFPASLQRVEI 283

Qy        276 ---------KKKNITKTF-KINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT 325
                        ||::||:|| ||:          :|    :|   |  | ::    | ::|:|
Db        284 KDGKGVACLKKEHITQTFPKIHDLVKQSI----FVSVSVLTEGGGEMVEAEKRG-IQIVT 338
```

```
Qy        326 TVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEER 385
Db        339 S--------------------PYSILFKRTPKYFKPGMPFDVSVYITNPD--------- 368

Qy        386 RNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKA 445
Db        369 -NSPAIGVEVEVTPDHAKGVTRAN---------------GFAKIPLNTVASATELVITV 411

Qy        446 ---------YFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVS-GNKRL-- 493
Db        412 KTKDPGDPRQQTGGGTMKALP--YRTSTKNFLHVGVDSNELKIGDPIKIDLNLGPTTIPN 469

Qy        494 KELSYMVVSRGQLVAVG---KQNSTMFSLT--PENSWTPKACVIVYYIEDDGEIISDVLK 548
Db        470 HDLTYMFLSRGQLVKVGRFKRQGNALVTLSVPVSKELLPSFRIVAYYHVGAADLVADSVW 529

Qy        549 IPVQLVFKNKIKLYWSKVKA--EPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDIT 606
Db        530 VDIKVSCMGSLKVTSTRPKASYEPRRAFSLTIT-GDPGAKVGLVAVDKGVYVLNSKHRLT 588

Qy        607 MENVVHELELYNTGYYLGMFMNSFAVFQECGL----------WVLTDAN---------- 645
Db        589 QTKIWDTIEKHDTGCTAGGGADNMGVFYDAGLVFETNTAKGTGIRTDPSCPVSSRRRAV 648

Qy        646 ---------------LTKD-YIDGVYDN--------------------------AEYAE 662
Db        649 TISDVITSMASKYHGLAKECCVDGMRDNTMGYTCDRRAQYISDGDVCVQAFLVCCTEMAS 708

Qy        663 RFMEENEGHIV-------DIHDFSLGSSPHV-RKHFPETWIWLDTNMGYRIYQ--EFEVT 712
Db        709 KKIESKQDALLLSRSEEDDDDAYMRSEDIVSRSQFPESWMWEDTNLPECPAQNKHCEST 768

Qy        713 -------VPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFA 765
Db        769 SVIRNNFLKDSITTWQITAISLSKTHGI-CVADPFEMIVLKEFFIDLKLPYSAVRNEQLE 827

Qy        766 LEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPT 825
Db        828 VKAILHNYSEDPIIVRVELMENGE----VCSSASKKGKYRQEVNMDPMSTRVVPYVIIPM 883

Qy        826 HLG--EIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFS 883
Db        884 KLGLHSIEVKASVKNSGSNDGVKRDLRVVAEGVLVKKETNVLLNPVKHGGEQT-SHIPSG 942

Qy        884 FPPNTVTGSER---VQITAIGD----VLGPSING--LASLIRMPYGCGEQNMINFAPNIY 934
Db        943 VPRNQVPNSDADTLISVTA-GEQTSVLVEQAISGDSLGSLIVQPVGCGEQNMIYMTLPVI 1001

Qy        935 ILDYLTKKKQLTD---NLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAF 991
Db       1002 ATHYLDNTKKWEDIGLDKRNTAIKYINIGYQRQLAYRKEDGSYAAW--VSRQSSTWLTAY 1059

Qy        992 VLRCFLEADPYIDIDQNVLHRTYTWL-KGHQKSNGEFWDPGRVIHSELQG---GNKSPVT 1047
Db       1060 VVKVFAMSSTLISVQENVLCTAVKWLILNTQQPDGIFNEFAPVIHAEMTGNVRGSDNDAS 1119

Qy       1048 LTAYI-------------VTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALIT 1093
Db       1120 MTAFVLIAMQEASSVCEQSVNSLPG--------SMAKAVAYLEKRLPH-LTNPYAVAMTS 1170

Qy       1094 YALSSVGSPKAKEALNMLTWRAEQ---EGGMQFWVSSESKLSDSWQPRSLDIEVAAYALL 1150
Db       1171 YALANAGKLNKETLLKFASPQLDHWPVPGGYQY-----------------TLEATSYALL 1213

Qy       1151 SHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTER---TNIQV 1207
Db       1214 ALVKVKAFEEAGPIVRWLNKQKKVGGGYGSTQSTIMVFQAVAEYWSHVKDLKDFDLNINL 1273

Qy       1208 TVTGPSSPSPLAVVQPMAVNISANGFGFAICQLNVV---YNVKASGSSRRRRSI------ 1258
Db       1274 EVAGRAS------VTKWSIN-NKNQFHTRTDKVNSIDKDLTVKASGNGEATLSVVTLYYA 1326

Qy       1259 ------QNQEAFDLDVAVKENKDDLNHVD------LNVCTSFSGPGRSG-MALMEVNLLS 1305
Db       1327 LPEEKDSDCESFDLSVTLTK-MDKTSHEDAKESFMLTIEVLYKNSERDATMSILDIGLLT 1385

Qy       1306 GFMVPSEAIS-LSE----TVKKVEYD-----HGKLNLYLDSVN---ETQFCVNIPAVRNF 1352
Db       1386 GFIVDTDDLNQLSKGRERYIEKFEMDKVLSERGSLILYLDKVSHKLEDRISFKIHRVQ-- 1443

Qy       1353 KVSNTQDASVSIVDYYEPRRQAVRSYNSE------VKLSSCDLCS 1391
Db       1444 EVGVLQPAAVSVYEYYNQKR-CVKFYHPQREGGTLSRLCLGDVCT 1487

RESULT 13
CO3_NAJNA
ID   CO3_NAJNA      STANDARD;      PRT;  1651 AA.
AC   Q01833;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C3 precursor [Contains: C3A anaphylatoxin].
GN   C3.
OS   Naja naja (Indian cobra).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Elapinae; Naja.
OX   NCBI_TaxID=35670;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=93056528; PubMed=1431125;
RA   Fritzinger D.C., Connelly M., Petrella E.C., Bredehorst R.,
RA   Vogel C.W.;
RT   "Primary structure of cobra complement component C3.";
RL   J. Immunol. 149:3554-3562(1992).
CC   -!- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
CC       REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.
CC       AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE
CC       THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,
CC       C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES.
CC   -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC       residues, forming two chains, beta and alpha, linked by a
CC       disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC       chain, releasing C3A anaphylatoxin and generating C3B (beta chain
CC       + alpha' chain) (By similarity).
CC   -!- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L02365; AAA49385.1; -.
DR   HSSP; P01024; 1C3D.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001840; Anaphylatoxn.
DR   InterPro; IPR001599; Macroglobln A2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   PRINTS; PR00004; ANAPHYLATOXN.
DR   ProDom; PD003264; Anaphylatoxin; 1.
```

```
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Complement alternate pathway; Plasma;
KW   Inflammatory response; Signal; Thioester bond.
FT   SIGNAL        1     22
FT   CHAIN        23   1651       COMPLEMENT C3.
FT   CHAIN        23    655       BETA CHAIN.
FT   CHAIN       661   1651       ALPHA CHAIN.
FT   PEPTIDE     661    738       C3A ANAPHYLATOXIN.
FT   CHAIN       739   1651       C3B (ALPHA' CHAIN).
FT   SITE        738    739       CLEAVAGE (BY C3 CONVERTASE).
FT   DOMAIN      683    718       ANAPHYLATOXIN-LIKE.
FT   DISULFID    546    807       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    615    650       BY SIMILARITY.
FT   DISULFID    683    710       BY SIMILARITY.
FT   DISULFID    684    717       BY SIMILARITY.
FT   DISULFID    697    718       BY SIMILARITY.
FT   DISULFID    863   1501       BY SIMILARITY.
FT   DISULFID   1091   1147       BY SIMILARITY.
FT   DISULFID   1346   1477       BY SIMILARITY.
FT   DISULFID   1377   1446       BY SIMILARITY.
FT   DISULFID   1494   1499       BY SIMILARITY.
FT   DISULFID   1506   1578       BY SIMILARITY.
FT   DISULFID   1525   1649       BY SIMILARITY.
FT   CROSSLNK    999   1002       Isoglutamyl cysteine thioester (Cys-Gln)
FT                                (By similarity).
SQ   SEQUENCE   1651 AA;  184926 MW;  66F91BD8F4A935F7 CRC64;

  Query Match             13.2%;  Score 973.5;  DB 1;  Length 1651;
  Best Local Similarity   23.5%;  Pred. No. 2.5e-45;
  Matches  391;  Conservative  278;  Mismatches  649;  Indels  343;  Gaps   61;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELL-EHCPSQVTVK 59
              |:|  |    | ||      :    :  |   :  :  | ::|        | ||    :   |   :  :
Db          1 MEGMALYLVAALLIGFPGS---SHGALYTLITPAVLRTDTEEQILVEAHGDSTPKSLDIF 57

Qy         60 A----ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEIL 115
                        ||         |  :  :|   :|           |: :|:  ||     :     |     |    ::
Db         58 VHDFPRKQKTLFQSRVDMNQAGSMFVT---PTIKVPAKELNKDSKQNQYVVVKVTGPQVA 114

Qy        116 FSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNL 175
                      ||::::      ||||||  :|  |         |:  :|:  :     |        ::       ::::
Db        115 LEKVVLLSYQSG--FVFIQTDKGIYTPGSPVRYRVFSV--DHNMHRMDKTVIVE------ 164

Qy        176 IQQWLSQQSDLGVISKTFQLSSHPI---------------LGDWSIQVQVN---DQTYY 216
                          |:   |::      :||  |:                      ||       :          :::|
Db        165 -----FQTPEGIV-----VSSKPVNPSGSIRPYNLPELVSFGTWKAVAKYEHSPEESYT 213

Qy        217 QSFQVSEYVLPKFEVTLQTP---LYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSF 273
                   | |  |||||  ||| ||          ||       |  |: :  :|||:|  |||  |:|  |             :
Db        214 AYFDVREYVLPSFEVRLQPSDKFLYIDGN-KNFHVSITARYLYGKKVEG-VAFVVFGVKI 271

Qy        274 WGKKKNI----TKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTE 329
                    ||:|        |:      |:|          :       : :::         |  |      :         :|  :||
Db        272 DDAKKSIPDSLTRIPIIDGDGEATLKRDTLRSRFQDLNQLVGHTLYVS------VTVITE 325

Qy        330 SVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNV 389
              |  :   :                    |  ||       ||       ||   :     ||     ||    ||:
Db        326 SGSDMVVTEQGGIHIVTSPYQIYFTKTPKYFKPGMPYELTVYVTNPDGSPAA-------H 378

Qy        390 VITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVP---QSGTFKIEFPILEDSSELQLKAY 446
               |  |::  ::|            |      :        |:        |       ||           ||       ::      |
Db        379 VPVVSEAIHSE-------GTTLSDGTAKLILNTPLNIQS------LPITVRTNHGDLPRE 425

Qy        447 FLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS--PFELVVSGN----KRLKELSYMV 500
                           ||  |            :  |:  |:  :           ||  |         |           |  ||           ::|    :|::
Db        426 RQAIKSMTATAYQTQGGSENYLHVAITSTEIKPGDNLPVNFNVRGNANSLNQIKYFTYLI 485

Qy        501 VSRGQLVAVGK------QNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQ-- 552
              :::|::   ||:          ||       :|          |         :  ||     :  ||::|  :  :  |:
Db        486 LNKGKIFKVGRQPRRDGQNLVTMNLHITPDLIPSFRFVAYYQVGNNEIVADSVWVDVKDT 545

Qy        553 ----LVFK------NKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNAS 602
                   || |          ::|:           :|    : :::       |  :  ||:|||||:|  ::|
Db        546 CMGTLVVKGASSRDDRIQ--------KPGAAMKIKLE-GDPGARVGLVAVDKAVYVLNDK 596

Qy        603 NDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLT--------------- 647
                 |:    :      :|    :   |      |      |:    ||::   ||   :   |    ||
Db        597 YKISQAKIWDTIEKSDFGCTAGSGQNNLGVFEDAGLALTTSTNLNTKQRSAAKCPQPANR 656

Qy        648 -------------------------KDYIDGVYDN-----AEYAERFMEENEG------- 670
                                                 |     ||::::|         |             ::::|   :
Db        657 RRSSVLLLDSKASKAAQFQDQGLRKCCEDGMHENPMGYTCEKRAKYIQEGDACKAAFLE 716

Qy        671 --HIV-----------------DIHDFSLGSSPHV-RKHFPETWIWL-------DTNMG 702
                 |  :                             |      |       |         :    |     |||:|:||                   |  |
Db        717 CCHYIKGIRDENQRESELFLARSDFEDELFGDDNIISRSDFPESWLWLTEELTGEPNNQG 776

Qy        703 YRIYQEFEVTVP----DSITSWVATGFVISEDLGLGLTTT-------PVELQAFQPFFIF 751
                              |||          ||||:|                  |  |  :||:  |                  |  |:      :  |||
Db        777 IS-----SKTVPFYLRDSITTW--------ELLAVGLSPTKGICVAEPYEITVMKDFFID 823

Qy        752 LNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVP 811
              |  ||||::  |:   :      ::||    :       |:|  :      :      |             |:          ::|        :
Db        824 LRLPYSVVKNEQVEIRAILYNYADEDIYVRVELIYNPAF----CSASTEGQRYRQQFPIK 879

Qy        812 SEDGATVLFPIRPTHLG--EIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLT 869
                 :        |    |     |     |      :: :   :              |||  |  :   :      ||     |:           |  ||    :
Db        880 ALSSRAVPFVIVPLEQGLHDVEVIASVRGELASDGVRKKLKVVPEGERKNIVTIIELDPS 939

Qy        870 DNRLQST--LKTLSFSFPPNTVTGSERVQITAIGD----VLGPSING--LASLIRMPYGC 921
                 :          |       |         ::                           :|:  :||        ::        ||:|    |       ||      |      ||
Db        940 VKGVGGTQELTVIANKLDDKVPDTEVETRISVLGDPVAQIIENSIDGSKLNHLIITPSGC 999

Qy        922 GEQNMINFAPNI---YILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFG 978
              |||||       |::         |  ||            :  :      |     :  :  :|:      :       ||   ::::|::  |  |::||
Db       1000 GEQNMITMTPSVIATYYLDATGQWENLGVDRRTEAIKQIMTGYAQQMVYKKADHSYAAFT 1059

Qy        979 NYDPSGSTWLSAFVLRCFLEADPYI-DIDQNVLHRTYTWL-KGHQKSNGEFWDPGRVIHS 1036
              |        :   |:||:|:|::              |              :  ||        ::                ||                |:  :|  |        :          |||
Db       1060 N--RASSSWLTAYVVKVLAMASNMVKDISHEIICGGVKWLILNRQQPDGVFKENAPVIHG 1117

Qy       1037 ELQGGNKS---PVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRG----------I 1083
              |:   ||  |            :|||:|||:||     |                  :|     |:     |:|         ::                            :
Db       1118 EMLGGTKGAEPEASLTAFIVTALLESRSV-----CKEQINILDSSINKATDYLLKKYEKL 1172

Qy       1084 SDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIE 1143
                           ||  ||     |||:                 |        :  ||                             :    :    :   |:||
Db       1173 QRPYTTALTAYALA------AADRLNDDRVLMAASTGRNRW--------EEYNARTHNIE 1218

Qy       1144 VAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTER- 1202
                  :||||:          :   :|      |::|||       |:              |          :        ||   |         :||:|:          |     |     :
Db       1219 GTSYALLALLKMKKFAEVGPVVRWLIDQKYYGGTYGQTQATVMVFQALAEYEIQMPTHQD 1278

Qy       1203 TNIQVTVTGPSSPSP-------LAVVQPMAV--------NISANGFGFAICQLNVVYNVK 1247
                |:   :::     |          |                              ||       |                       :||:|    |      |             :     |||   :
Db       1279 LNLDISIKLPEREVPERYSINDRNAVQARTVETKLNEDFTVSASGDGKATMTILTVYNAQ 1338

Qy       1248 ASGSSRRRRSIQNQEAFDLDVAV-------KENKDDLNHVDLNVCTSFSGPGRSGMALME 1300
                                |          ::     |:     |     |||:|                    |:    |                  :     |     :|| :    |                ||    |    :::
Db       1339 L----REDANVCNK--FHLDVSVENVELNLKQAKGGKAALRLKICTRYLGEVDSTMTIID 1392

Qy       1301 VNLLSGFMVPSEAI-----SLSETVKKVEYDH-----GKLNLYLYLDSVNETQ-FCVNIPAV 1349
              :::|:|     |   :|     |:                             :         |     :    :      |:      |:  |      :  |         ||:   ::     ||  |:  :      |   :      |:    |::
Db       1393 ISMLTGFFPDAEDLKRLSNGVDRYISKFEIDNNMAQKGTVVIYLDKVSHSEDECLHFKIH 1452

Qy       1350 RNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHH 1409
              ::|:|          |        ||      :  |                |   |          |   :                   :   :      ||      |       :
Db       1453 KHFEVGFIQPGSVKVYSYYNLDEQCTKFYHPDKETGLLNKICHGNICRCAEETCSLLNQQ 1512

Qy       1410 --------------SSVIFIFCFKLL----------YFMEL 1426
                                       :|  :::      |||                      |||::
Db       1513 KKIDLQLRIQKACAQNVDYVYKTKLLRIEEKDGNDIYFMDV 1553
```

```
RESULT 14
CO3_RAT
ID   CO3_RAT        STANDARD;      PRT;  1663 AA.
AC   P01026;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C3 precursor [Contains: C3A anaphylatoxin].
GN   C3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar; TISSUE=Liver;
RX   MEDLINE=90245672; PubMed=2336397;
RA   Misumi Y., Sohda M., Ikehara Y.;
RT   "Nucleotide and deduced amino acid sequence of rat complement C3.";
RL   Nucleic Acids Res. 18:2178-2178(1990).
RN   [2]
RP   SEQUENCE OF 671-748.
RX   MEDLINE=79062262; PubMed=309768;
RA   Jacobs J.W., Rubin J.S., Hugli T.E., Bogardt R.A., Mariz I.K.,
RA   Daniels J.S., Daughaday W.H., Bradshaw R.A.;
RT   "Purification, characterization, and amino acid sequence of rat
RT   anaphylatoxin (C3a).";
RL   Biochemistry 17:5031-5038(1978).
RN   [3]
RP   SEQUENCE OF 1316-1595 FROM N.A.
RX   MEDLINE=89380332; PubMed=2674144;
RA   Sundstrom S.A., Komm B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,
RA   Lyttle C.R.;
RT   "Estrogen regulation of tissue-specific expression of complement C3.";
RL   J. Biol. Chem. 264:16941-16947(1989).
CC   -!- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
CC       REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.
CC       AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE
CC       THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,
CC       C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES.
CC   -!- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG
CC       RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE
CC       BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,
CC       RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA'
CC       CHAIN).
CC   -!- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X52477; CAA36716.1; -.
DR   EMBL; M29866; AAA40837.1; ALT_SEQ.
DR   PIR; S15764; C3RT.
DR   PDB; 1QQF; 31-JUL-00.
DR   PDB; 1QSJ; 31-JUL-00.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; Macroglob1nA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Complement alternate pathway; Plasma;
KW   Inflammatory response; Glycoprotein; Signal; 3D-structure;
KW   Thioester bond.
FT   SIGNAL        1     24
FT   CHAIN        25   1663       COMPLEMENT C3.
FT   CHAIN        25    666       BETA CHAIN.
FT   CHAIN       671   1663       ALPHA CHAIN.
FT   PEPTIDE     671    748       C3A ANAPHYLATOXIN.
FT   CHAIN       749   1663       C3B (ALPHA' CHAIN).
FT   SITE        748    749       CLEAVAGE (BY C3 CONVERTASE).
FT   DOMAIN      693    728       ANAPHYLATOXIN-LIKE.
FT   DISULFID    558    816       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    626    661       BY SIMILARITY.
FT   DISULFID    693    720       BY SIMILARITY.
FT   DISULFID    694    727       BY SIMILARITY.
FT   DISULFID    707    728       BY SIMILARITY.
FT   DISULFID    873   1513       BY SIMILARITY.
FT   DISULFID   1101   1158       BY SIMILARITY.
FT   DISULFID   1358   1489       BY SIMILARITY.
FT   DISULFID   1389   1458       BY SIMILARITY.
FT   DISULFID   1506   1511       BY SIMILARITY.
FT   DISULFID   1518   1590       BY SIMILARITY.
FT   DISULFID   1537   1661       BY SIMILARITY.
FT   CROSSLNK   1010   1013       Isoglutamyl cysteine thioester (Cys-Gln).
FT   CARBOHYD    939    939       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   CARBOHYD   1617   1617       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   CONFLICT    721    722       LK -> KL (IN REF. 2).
SQ   SEQUENCE   1663 AA;  186460 MW;  2F87CCB143CDD4BC CRC64;

  Query Match             13.1%;  Score 966;  DB 1;  Length 1663;
  Best Local Similarity   24.3%;  Pred. No. 6.4e-45;
  Matches  391;  Conservative  284;  Mismatches  609;  Indels  326;  Gaps   68;

Qy          3 GPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLE---HCPSQVTVK 59
              |  ||        ||  :  ::|      |  :  :  ::|      |  :|   :      |  |||:
Db          6 GSQLLV---LLLLLASSLLALGSPMYSIITPNVLRLESEETFILEAHDAQGDVPVTVTVQ 62

Qy         60 AELLKTASNLTVSVLE-AEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSN 118
                          :||  |  |      |     :|:    :||: ::               :
Db         63 DFLKKQVLTSEKTVLTGATGHLNRVFIK---IPASKEFNADKGHKYVTVVANFGATVVEK 119

Qy        119 STRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQ 178
              :  :||::  :  |||||| :| |    | :||  |: ::   |    :: |:|: |    |::
Db        120 AVLVSFQSGYL--FIQTDKTIYTPGSTVFYRIFTVDNNLLPVGKTVVIVIETPDGVPIKR 177

Qy        179 -WLSQQSDLGVISKTFQLSSHPILGDWSIQV---QVNDQTYYQSFQVSEYVLPKFEVTLQ 234
                || :  |::  :: :     :| | |:  .      ||:    |:| ||||| ||| ::
Db        178 DILSSHNQYGILPLSWNIPELVNMGQWKIRAFYEHAPKQTFSAEFEVKEYVLPSFEVLVE 237

Qy        235 --TPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTF----------LPLSFWGKKKNITK 282
                  |        | |   :|||:: |||  |   |   :  |              | ||         :|:
Db        238 PTEKFYYIHGPKGLEVSITARFLYGKNVDGTAFVIFGVQDEDKKISLALS-------LTR 290

Qy        283 TFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNV 342
                   :||     :     | :||         :  |||               |:: |  ||  |
Db        291 VLIEDGSGEAVLS---RKVLMDG-------VRPSSP-----------EALVGKSLYVSVTV 330

Qy        343 FFKQ---------------HDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERR 386
                                  | | |         ||:: |    | ||  ||:            ||
Db        331 ILHSGSDMVEAERSGIPIVTSPYQIHFTKTPKFFKPAMPFDLMVFVTNPDGS----PARR 386

Qy        387 NNVVITVTQRNYTEYWSGSNS-GNQKMEAVQKINYTVPQS--------GTFKIEFPILED 437
                 |  |||        ||::    : :  | |::    |  :               |  |    |
Db        387 ---VPVVTQ--------GSDAQALTQDDGVAKLSVNTPNNRQPLTITVSTKKEGIPDARQ 435
```

```
Qy        438 SSE-LQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKEL 496
              :: :| : |     |: :|     | : ::|| | |: |             :::
Db        436 ATRTMQAQPYSTMHNSNNYLHL-----SVSRVELKPGD-NLNVNFHLRTDAGQEAKIRYY 489

Qy        497 SYMVVSRGQLVAVGK------QNSTMFSL--TPENSWTPKACVIVYY--IEDDG--EIIS 544
              :|:|:::|:|: |:        |: : || ||| : |  :: ||  | :| |:::
Db        490 TYLVMNKGKLLKAGRQVREPGQDLVVLSLPITPE--FIPSFRLVAYYTLIGANGQREVVA 547

Qy        545 DVLKIPVQ------LVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNL 598
              | : : |:      || |   :     :  |  : :|||  |  : ||:||||| | :
Db        548 DSVWVDVKDSCVGTLVVKGDPR---DNRQPAPGHQTTLRIEGNQ-GARVGLVAVDKGVFV 603

Qy        599 MNASNDITMENVVHELELYNTG-------YYLGMFMNSFAVFQ------------ECG- 637
              :|  | :|   :  :| : |         | |:||::   |:             ||
Db        604 LNKKNKLTQSKIWDVVEKADIGCTPGSGKNYAGVFMDAGLTFKTNQGLQTDQREDPECAK 663

Qy        638 -----------------LWVLTDANLTKDYIDGV------------------------ 654
                                  || | |   ||:
Db        664 PAARRRRSVQLMERRMDKAGQYTDKGLRKCCEDGMRDIPMPYSCQRRARLITQGESCLKA 723

Qy        655 -YDNAEYAERFMEEN-EGHIV-----DIHDFSLGSSPHV-RKHFPETWIW---------- 696
                |   |  : |:: |::      |: :  :    : | ||||:|:|
Db        724 FMDCCNYITKLREQHRRDHVLGLARSDVDEDIIPEEDIISRSHFPESWLWTIEELKEPEK 783

Qy        697 --LDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNL 754
                : | :        : : ||||:|     :|: |:      | |:  | ||| | |
Db        784 NGISTKV-------MNIFLKDSITTWEILAVSLSDKKGI-CVADPYEITVMQDFFIDLRL 835

Qy        755 PYSVIRGEEFALEITIFNYLKDATEVKVIIE--KSDKFDILMTSSEINATGHQQTLLVPS 812
              ||||:| |:  :     :||| :|    :  |  : |:  :       : ||: :|
Db        836 PYSVVRNEQVEIRAVLFNY-REQEKLKVRVELLHNPAFCSMATAKK----RYYQTIEIPP 890

Qy        813 EDGATVLFPIRPTHLG--EIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSI-LLD-- 867
              :     | : | |  |  |:  :        || | ::: |  ||: : : :: ||
Db        891 KSSVAVPYVIVPLKIGLQEVEVKAAVFNHFISDGVKKILKVVPEGMRVNKTVAVRTLDPE 950

Qy        868 -LTDNRLQS---TLKTLSFSFPPNTVTGSER---VQITAIGDVLGPSING--LASLIRMP 918
               |    :|        || |   | ||     :| | :  :  :::| | || |
Db        951 HLNQGGVQREDVNAADLSDQVPD---TDSETRILLQGTPVAQMAEDAVDGERLKHLIVTP 1007

Qy        919 YGCGEQNMINFAPNIYILDYLTKKKQLTD---NLKEKALSFMRQGYQRELLYQREDGSFS 975
               |||||||  | :  : || : :|        :::||  :::|| ::| :::   :::
Db       1008 SGCGEQNMIGMTPTVIAVHYLDQTEQWEKFGLEKRQEALELIKKGYTQQLAFKQPISAYA 1067

Qy        976 AFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWL-KGHQKSNGEFWDPGRVI 1034
              || |  | ||||:|    | |   | ||  ||        ||      || :| | : | ||
Db       1068 AFNNRPP--STWLTAMWSRSFSLAANLIAIDSQVLCGAVKWLILEKQKPDGVFQEDGPVI 1125

Qy       1035 HSELQGG----NKSPVTLTAYIVTSLLGYR---KYQPNI---DVQESIHFLESEFSRGIS 1084
               | |: ||      :: :|:|||::: :|  :|       :: ||| ||: :  :
Db       1126 HQEMIGGFRNTKEADVSLTAFVLIALQEARDICEGQVNSLPGSINKAGEYLEASY-LNLQ 1184

Qy       1085 DNYTLALITYAL---SSVGSPKAKEALNMLTWRAE-QEGGMQFWVSSESKLSDSWQPRSL 1140
                ||:|: |||    : :  |   : ||      :| | | :
Db       1185 RPYTVAIAGYALALMNKLEEPYLTKFLNTAKDRNRWEEPGQQLY--------------- 1228

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAA-LMN 1199
               ::|  :||||: |         |::|||: :|  ||: ||| | : :||::: | : :
Db       1229 NVEATSYALLALLLLKDFDSVPPVVRWLNDERYYGGGYGSTQATFMVFQALAQYRADVPD 1288

Qy       1200 TERTNIQVTVTGPSSPSPLA---------------VVQPMAVNISANGFGFAICQLNVVY 1244
               : |: |:: || ||                 |  :::| | |     : ||
Db       1289 HKDLNMDVSLHLPSRSSPTVFRLLWESGSLLRSEETKQNEGFSLTAKGKGQGTLSVVTVY 1348

Qy       1245 NVKASGSSRRRRSIQNQEAFDLDVAV-------KENKDDLNHVDLNVCTSFSGPGRSGMA 1297
              : | | :  ::         ||| |        |: :|   : : |::|| : |  | : |:
Db       1349 HAKVKGKTTCKK-------FDLRVTIKPAPETAKKPQDAKSSMILDICTRYLGDVDATMS 1401

Qy       1298 LMEVNLLSGFMVPSEAISLSET-----VKKVEYDHGKLN-----LYLDSVNET-QFCVNI 1346
              ::::::::||: :  : | :       : | | |    |      :||: :: : : |::
Db       1402 ILDISMMTGFIPDTNDLELLSSGVDRYISKYEMDKAFSNKNTLIIYLEKISHSEEDCLSF 1461

Qy       1347 PAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSE------VKLSSCDLC 1390
                : | |   | || : ||           | |: |       ||   ::|
Db       1462 KVHQFFNVGLIQPGSVKVYSYYNLEESCTRFYHPEKDDGMLSKLCHNEMC 1511
```

```
RESULT 15
CO3_MOUSE
ID   CO3_MOUSE      STANDARD;      PRT;  1663 AA.
AC   P01027;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Complement C3 precursor (HSE-MSF) [Contains: C3A anaphylatoxin].
GN   C3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM LONG).
RX   MEDLINE=85038854; PubMed=6208565;
RA   Fey G.H., Lundwall A., Wetsel R.A., Tack B.F., de Bruijn M.H.L.,
RA   Domdey H.;
RT   "Nucleotide sequence of complementary DNA and derived amino acid
RT   sequence of murine complement protein C3.";
RL   Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:333-344(1984).
RN   [2]
RP   SEQUENCE OF 671-1663 FROM N.A. (ISOFORM LONG).
RX   MEDLINE=85054819; PubMed=6094532;
RA   Wetsel R.A., Lundwall A., Davidson F., Gibson T., Tack B.F., Fey G.H.;
RT   "Structure of murine complement component C3. II. Nucleotide sequence
RT   of cloned complementary DNA coding for the alpha chain.";
RL   J. Biol. Chem. 259:13857-13862(1984).
RN   [3]
RP   SEQUENCE OF 671-748 FROM N.A.
RX   MEDLINE=83117730; PubMed=6961437;
RA   Domdey H., Wiebauer K., Kazmaier M., Mueller V., Odink K., Fey G.H.;
RT   "Characterization of the mRNA and cloned cDNA specifying the third
RT   component of mouse complement.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:7619-7623(1982).
RN   [4]
RP   SEQUENCE OF 658-761 FROM N.A.
RX   MEDLINE=84201365; PubMed=6609661;
RA   Fey G.H., Wiebauer K., Domdey H.;
RT   "Amino acid sequences of mouse complement C3 derived from nucleotide
RT   sequences of cloned cDNA.";
RL   Ann. N.Y. Acad. Sci. 421:307-312(1983).
RN   [5]
RP   SEQUENCE OF 1-34 FROM N.A.
RX   MEDLINE=83117622; PubMed=6985486;
RA   Wiebauer K., Domdey H., Diggelmann H., Fey G.;
RT   "Isolation and analysis of genomic DNA clones encoding the third
RT   component of mouse complement.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:7077-7081(1982).
RN   [6]
RP   SEQUENCE OF 25-41 AND 749-760.
RX   MEDLINE=93373334; PubMed=8364938;
RA   Hamada J.-I., Cavanaugh P.G., Miki K., Nicolson G.L.;
RT   "A paracrine migration-stimulating factor for metastatic tumor cells
RT   secreted by mouse hepatic sinusoidal endothelial cells:
RT   identification as complement component C3b.";
RL   Cancer Res. 53:4418-4423(1993).
RN   [7]
RP   ALTERNATIVE INITIATION.
RX   MEDLINE=95053742; PubMed=7964485;
RA   Cahen-Kramer Y., Martensson I.L., Melchers F.;
RT   "The structure of an alternate form of complement C3 that displays
RT   costimulatory growth factor activity for B lymphocytes.";
RL   J. Exp. Med. 180:2079-2088(1994).
CC   -!- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC       COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
CC       REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.
CC       AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE
```

```
CC       THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.
CC   -!- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,
CC       C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC       INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC       PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC       BASOPHILIC LEUKOCYTES. THE SHORT ISOFORM HAS B-CELL STIMULATORY
CC       ACTIVITY.
CC   -!- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC       residues, forming two chains, beta and alpha, linked by a
CC       disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC       chain, releasing C3A anaphylatoxin and generating C3B (beta chain
CC       + alpha' chain).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative initiation;
CC         Comment=2 isoforms, Long (shown here) and Short, are produced by
CC         alternative initiation;
CC   -!- MISCELLANEOUS: C3B IS RAPIDLY SPLIT IN TWO POSITIONS BY FACTOR I
CC       AND A COFACTOR TO FORM IC3B (INACTIVATED C3B) AND C3F WHICH IS
CC       RELEASED.
CC   -!- MISCELLANEOUS: IC3B IS THE SLOWLY CLEAVED (POSSIBLY BY FACTOR I)
CC       TO FORM C3C AND C3DG. OTHER PROTEASES PRODUCE OTHER FRAGMENTS SUCH
CC       AS C3D OR C3G.
CC   -!- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC   -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; K02782; AAC42013.1; -.
DR   EMBL; J00369; AAA37336.1; -.
DR   EMBL; J00367; AAA37336.1; JOINED.
DR   EMBL; M33032; AAA37378.1; -.
DR   EMBL; Z37998; CAA86099.2; -.
DR   PIR; A92459; C3MS.
DR   HSSP; P01024; 1C3D.
DR   MGD; MGI:88227; C3.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000020; Anaphylatoxin.
DR   InterPro; IPR001599; MacroglobinA2.
DR   InterPro; IPR001134; Netrin_C.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF01821; ANATO; 1.
DR   Pfam; PF01759; NTR; 1.
DR   ProDom; PD003264; Anaphylatoxin; 1.
DR   SMART; SM00104; ANATO; 1.
DR   SMART; SM00643; C345C; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
DR   PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
KW   Complement pathway; Complement alternate pathway; Plasma;
KW   Inflammatory response; Glycoprotein; Signal; Alternative initiation;
KW   Thioester bond.
FT   SIGNAL        1     24
FT   CHAIN        25   1663       COMPLEMENT C3, ISOFORM LONG.
FT   CHAIN        25    666       COMPLEMENT C3, BETA CHAIN.
FT   CHAIN       671   1663       COMPLEMENT C3, ALPHA CHAIN.
FT   CHAIN      1129   1663       COMPLEMENT C3, ISOFORM SHORT.
FT   INIT_MET   1129   1129       FOR ISOFORM SHORT.
FT   PEPTIDE     671    748       C3A ANAPHYLATOXIN.
FT   CHAIN       749   1663       C3B (ALPHA' CHAIN).
FT   PEPTIDE     749    954       C3C FRAGMENT.
FT   PEPTIDE     955   1303       C3DG FRAGMENT.
FT   PEPTIDE     955   1001       C3G FRAGMENT.
FT   PEPTIDE    1002   1303       C3D FRAGMENT.
FT   PEPTIDE    1304   1320       C3F FRAGMENT.
FT   SITE        748    749       CLEAVAGE (BY C3 CONVERTASE).
FT   SITE       1303   1304       CLEAVAGE (BY FACTOR I).
FT   SITE       1320   1321       CLEAVAGE (BY FACTOR I).
FT   DOMAIN      693    728       ANAPHYLATOXIN-LIKE.
FT   DISULFID    559    816       INTERCHAIN (BY SIMILARITY).
FT   DISULFID    626    661       BY SIMILARITY.
FT   DISULFID    693    720       BY SIMILARITY.
FT   DISULFID    694    727       BY SIMILARITY.
FT   DISULFID    707    728       BY SIMILARITY.
FT   DISULFID    873   1513       BY SIMILARITY.
FT   DISULFID   1101   1158       BY SIMILARITY.
FT   DISULFID   1358   1489       BY SIMILARITY.
FT   DISULFID   1389   1458       BY SIMILARITY.
FT   DISULFID   1506   1511       BY SIMILARITY.
FT   DISULFID   1518   1590       BY SIMILARITY.
FT   DISULFID   1537   1661       BY SIMILARITY.
FT   DISULFID   1637   1646       BY SIMILARITY.
FT   CARBOHYD    939    939       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   1617   1617       N-LINKED (GLCNAC. . .).
FT   CROSSLNK   1010   1013       Isoglutamyl cysteine thioester (Cys-Gln)
FT                                (By similarity).
SQ   SEQUENCE   1663 AA;  186482 MW;  DE5546CC769BEA19 CRC64;

  Query Match             13.0%;  Score 956;  DB 1;  Length 1663;
  Best Local Similarity   24.0%;  Pred. No. 2.3e-44;
  Matches  384;  Conservative  286;  Mismatches  626;  Indels  302;  Gaps   66;

Qy          3 GPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLE---HCPSQVTVK 59
              |  ||    ||     ||  :   | : :  | ::|     || :|  :     |  |||:
Db          6 GSQLLVLLLLLASSPLALGI---PMYSIITPNVLRLESEETIVLEAHDAQGDIPVTVTVQ 62

Qy         60 AELLKTASNLTVSVLE-AEGVFEKGSFKTLTLP-SLPLNSADEIYELRVTGRTQDEILFS 117
                | :       :||  | |     ||    :| |   ||  | ::         | :
Db         63 DFLKRQVLTSEKTVLTGASGHLRSVSIK---IPASKEFNSDKEGHKYVTVVANFGETVVE 119

Qy        118 NSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQ 177
               :  :||::  :  |||||| :| |   | :|| |: ::  |   :: |||: |    ::
Db        120 KAVMVSFQSGYL--FIQTDKTIYTPGSTVLYRIFTVDNNLLPVGKTVVILIETPDGIPVK 177

Qy        178 Q-WLSQQSDLGVISKTFQLSSHPILGDWSIQV---QVNDQTYYQSFQVSEYVLPKFEVTL 233
              :  ||  :  |::  :: :     :| | |:        | :   |:| ||||| ||| :
Db        178 RDILSSNNQHGILPLSWNIPELVNMGQWKIRAFYEHAPKQIFSAEFEVKEYVLPSFEVRV 237

Qy        234 Q--TPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKK-----NITKTFKI 286
              :     |   :   |  :| ||: ||| | |   : |      | ||     ::|:
Db        238 EPTETFYYIDDPNGLEVSIIAKFLYGKNVDGTAFVIFGVQD--GDKKISLAHSLTRVVIE 295

Qy        287 NGSANFSFNDEEMKNVMDSSN-----GLSEYLDLSSPGPVEILTTVTESVTGISRNVSTN 341
              :|  :     : :   :  ||     | | |:          :| :  | : :     :
Db        296 DGVGDAVLTRKVLMEGVRPSNADALVGKSLYVS---------VTVILHSGSDMVEAERSG 346

Qy        342 VFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEY 401
              :      | | |       ||:: |   | ||  ||:         : |: |||
Db        347 IPIVTSPYQIHFTKTPKFFKPAMPFDLMVFVTNPDGSPA-------SKVLVVTQ------ 393

Qy        402 WSGSNS-GNQKMEAVQKINYTVPQSG-----TFKIEFPILEDSSE----LQLKAYFLGSK 451
                |||:    : : | |::   | |      | : :   | :| :    ::   |
Db        394 --GSNAKALTQDDGVAKLSINTPNSRQPLTITVRTKKDTLPESRQATKTMEAHPYSTMHN 451

Qy        452 SSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVAVGK 511
              |:  :|      | : ::||  | |: |            :::  :|:|:::|:|:  |:
Db        452 SNNYLHL-----SVSRMELKPGD-NLNVNFHLRTDPGHEAKIRYYTYLVMNKGKLLKAGR 505

Qy        512 ------QNSTMFSL--TPENSWTPKACVIVYY----IEDDGEIISDVLKIPVQ------L 553
                    |:  : ||  |||  : |   :: ||         |:::| : : |:      |
Db        506 QVREPGQDLVVLSLPITPE--FIPSFRLVAYYTLIGASGQREVVADSVWVDVKDSCIGTL 563

Qy        554 VFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHE 613
              | |   :         | :: :|||   |  : ||:||||| | ::|  | :|   :
Db        564 VVKGDPR----DNHLAPGQQTTLRIEGNQ-GARVGLVAVDKGVFVLNKKNKLTQSKIWDV 618

Qy        614 LELYNTG-------YYLGMFMNSFAVFQ-------------EC----------------- 636
              :|  : |        | |:||::   |:             ||
Db        619 VEKADIGCTPGSGKNYAGVFMDAGLAFKTSQGLQTEQRADLECTKPAARRRRSVQLMERR 678
```

```
Qy    637 ----GLWVLTDANLTKDYIDGVYD-----NAEYAERFMEENE----------GHIVDI-- 675
Db    679 MDKAGQY--TDKGLRKCCEDGMRDIPMRYSCQRRARLITQGENCIKAFIDCCNHITKLRE 736

Qy    676 -----HDFSLGSS-------PH----VRKHFPETWIW------------LDTNMGYRIYQ 707
Db    737 QHRRDHVLGLARSELEEDIIPEEDIISRSHFPQSWLWTIEELKEPEKNGISTKV------ 790

Qy    708 EFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALE 767
Db    791 -MNIFLKDSITTWEILAVSLSDKKGI-CVADPYEIRVMQDFFIDLRLPYSVVRNEQVEIR 848

Qy    768 ITIFNYLKDATEVKVIIE--KSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPT 825
Db    849 AVLFNY-REQEELKVRVELLHNPAFCSMATAK----NRYFQTIKIPPKSSVAVPYVIVPL 903

Qy    826 HLG--EIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTD---------NRLQ 874
Db    904 KIGQQEVEVKAAVFNHFISDGVKKTLKVVPEGMR--INKTVAIHTLDPEKLGQGGVQKVD 961

Qy    875 STLKTLSFSFPPNTVTGSER---VQITAIGDVLGPSING--LASLIRMPYGCGEQNMINF 929
Db    962 VPAADLSDQVPD---TDSETRIILQGSPVVQMAEDAVDGERLKHLIVTPAGCGEQNMIGM 1018

Qy    930 APNIYILDYLTKKKQLTD---NLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGST 986
Db   1019 TPTVIAVHYLDQTEQWEKFGIEKRQEALELIKKGYTQQLAFKQPSSAYAAFNNRPP--ST 1076

Qy    987 WLSAFVLRCFLEADPYIDIDQNVLHRTYTWL-KGHQKSNGEFWDPGRVIHSELQGG---- 1041
Db   1077 WLTAYVVKVFSLAANLIAIDSHVLCGAVKWLILEKQKPDGVFQEDGPVIHQEMIGGFRNA 1136

Qy   1042 NKSPVTLTAYIVTSLLGYR---KYQPNI---DVQESIHFLESEFSRGISDNYTLALITYA 1095
Db   1137 KEADVSLTAFVLIALQEARDICEGQVNSLPGSINKAGEYIEASY-MNLQRPYTVAIAGYA 1195

Qy   1096 L---SSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSH 1152
Db   1196 LALMNKLEEPYLGKFLNTAKDRNR-------WEEPDQQL--------YNVEATSYALLAL 1240

Qy   1153 FLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAA-LMNTERTNIQVTVTG 1211
Db   1241 LLLKDFDSVPPVVRWLNEQRYYGGGYGSTQATFMVFQALAQYQTDVPDHKDLNMDVSFHL 1300

Qy   1212 PSSPSPLA---------------VVQPMAVNISANGFGFAICQLNVVYNVKASGSSRRRR 1256
Db   1301 PSRSSATTFRLLWENGNLLRSEETKQNEAFSLTAKGKGRGTLSVVAVYHAKL-------K 1353

Qy   1257 SIQNQEAFDLDVAV-------KENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMV 1309
Db   1354 SKVTCKKFDLRVSIRPAPETAKKPEEAKNTMFLEICTKYLGDVDATMSILDISMMTGFAP 1413

Qy   1310 PSEAISL-----SETVKKVEYDHGKLN-----LYLDSVNET-QFCVNIPAVRNFKVSNTQ 1358
Db   1414 DTKDLELLASGVDRYISKYEMNKAFSNKNTLIIYLEKISHTEEDCLTFKVHQYFNVGLIQ 1473

Qy   1359 DASVSIVDYYEPRRQAVRSYNSE------VKLSSCDLC 1390
Db   1474 PGSVKVYSYYNLEESCTRFYHPEKDDGMLSKLCHSEMC 1511
```

Search completed: January 15, 2004, 18:12:59
Job time : 31 secs

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:         January 15, 2004, 18:09:19 ; Search time 53 Seconds
                                             (without alignments)
                                             6952.814 Million cell updates/sec

Title:          US-10-020-095-4
Perfect score:  7348
Sequence:       1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters:      830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_23:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_virus:*
                13:  sp_vertebrate:*
                14:  sp_unclassified:*
                15:  sp_rvirus:*
                16:  sp_bacteriap:*
                17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                     %
Result             Query
  No.    Score     Match Length DB  ID                          Description
-----------------------------------------------------------------------------
     1  7326.5      99.7    1445  4  Q8TDJ3                     Q8tdj3 homo sapien
     2    5360      72.9    1442 11  Q8R422                     Q8r422 mus musculu
     3  4305.5      58.6     854  4  Q8N3A7                     Q8n3a7 homo sapien
     4    3360      45.7     665  4  Q8N915                     Q8n915 homo sapien
     5    1850      25.2    1519  5  O46015                     O46015 caenorhabdi
     6  1779.5      24.2    1809  5  Q8T398                     Q8t398 ciona intes
     7    1703      23.2    1388  5  Q8IPH5                     Q8iph5 drosophila
     8  1699.5      23.1    1397  5  Q9VLZ0                     Q9vlz0 drosophila
     9    1693      23.0    1408  5  Q8IPH4                     Q8iph4 drosophila
    10    1692      23.0    1420  5  Q9NFV7                     Q9nfv7 drosophila
    11  1688.5      23.0    1399  5  Q9VLY9                     Q9vly9 drosophila
    12    1638      22.3    1507  5  O01717                     O01717 limulus sp.
    13  1542.5      21.0    1884  4  Q9ULD7                     Q9uld7 homo sapien
    14  1542.5      21.0    1885  4  Q8IZJ3                     Q8izj3 homo sapien
    15    1491      20.3    1476 11  Q60486                     Q60486 cavia porce
    16  1463.5      19.9    1496  5  Q9NFV5                     Q9nfv5 drosophila
    17  1461.5      19.9    1496  5  Q9VIT9                     Q9vit9 drosophila
    18    1460      19.9    1354  5  Q9NFV8                     Q9nfv8 drosophila
    19  1455.5      19.8    1494  5  Q8IT76                     Q8it76 ornithodoro
    20  1428.5      19.4    1732  5  Q969A4                     Q969a4 branchiosto
    21  1425.5      19.4    1469  5  Q9NFV6                     Q9nfv6 drosophila
    22  1425.5      19.4    1469  5  Q9VLY7                     Q9vly7 drosophila
    23    1424      19.4    1503 13  Q91076                     Q91076 lampetra ja
    24    1417      19.3    1487 11  Q03626                     Q03626 rattus norv
    25    1384      18.8    1461 13  O12978                     O12978 xenopus lae
    26    1384      18.8    1500 11  Q63041                     Q63041 rattus norv
    27    1384      18.8    1500 11  Q63332                     Q63332 rattus norv
    28  1374.5      18.7    1464 11  Q60488                     Q60488 cavia porce
    29    1354      18.4    1269  5  Q9NK75                     Q9nk75 drosophila
    30  1328.5      18.1    1760  5  Q9VLT3                     Q9vlt3 drosophila
    31    1321      18.0    1442 13  Q9PVU5                     Q9pvu5 cyprinus ca
    32  1310.5      17.8    1728  5  Q8I7P1                     Q8i7p1 swiftia exs
    33  1284.5      17.5    1406 13  Q9PVU4                     Q9pvu4 cyprinus ca
    34  1233.5      16.8    1340  5  Q9GYW4                     Q9gyw4 anopheles g
    35  1105.5      15.0    1699  5  O44344                     O44344 strongyloce
    36  1099.5      15.0     798  5  Q8MT94                     Q8mt94 drosophila
    37  1079.5      14.7    1683 13  Q91741                     Q91741 xenopus lae
    38    1073      14.6    1700 13  Q9I933                     Q9i933 cyprinus ca
    39  1054.5      14.4    1489 13  Q8QGD4                     Q8qgd4 rana catesb
    40    1054      14.3    1716 13  Q9I932                     Q9i932 cyprinus ca
    41     990      13.5    1652 13  Q90633                     Q90633 gallus gall
    42     985      13.4    1738 11  O70346                     O70346 mus musculu
    43     983      13.4    1614 13  Q98977                     Q98977 oncorhynchu
    44     965      13.1    1662 13  Q98TS6                     Q98ts6 anarhichas
    45     963      13.1     785 11  Q62591                     Q62591 rattus sord

                              ALIGNMENTS

RESULT 1
Q8TDJ3
ID   Q8TDJ3      PRELIMINARY;      PRT;  1445 AA.
AC   Q8TDJ3;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   CD109.
GN   CD109.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21849742; PubMed=11861284;
RA   Lin M., Sutherland D.R., Horsfall W., Totty N., Yeo E., Nayar R.,
RA   Wu X.F., Schuh A.C.;
RT   "Cell surface antigen CD109 is a novel member of the alpha(2)
RT   macroglobulin/C3, C4, C5 family of thioester-containing proteins.";
RL   Blood 99:1683-1691(2002).
DR   EMBL; AF410459; AAL84159.1; -.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
SQ   SEQUENCE   1445 AA;  161719 MW;  6EE6D2D936AF310D CRC64;

  Query Match             99.7%;  Score 7326.5;  DB 4;  Length 1445;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 1427;  Conservative    1;  Mismatches    0;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy       121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy       181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy       241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy       301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy       361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy       421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy       481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy       541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy       601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy       661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720

Qy       721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy       781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
             ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db       781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy       841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy       901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy       961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy      1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy      1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy      1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200

Qy      1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
             |||||||||||||||||                 ||||||||||||||||||||||||||
Db      1201 ERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVV 1260

Qy      1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1320

Qy      1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1321 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1380

Qy      1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHSSVIFIFCFKLLYF 1423
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1381 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHSSVIFIFCFKLLYF 1440

Qy      1424 MELWL 1428
             |||||
Db      1441 MELWL 1445
```

```
RESULT 2
Q8R422
ID   Q8R422      PRELIMINARY;      PRT;  1442 AA.
AC   Q8R422;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   GPI-anchored alpha-2 macroglobulin-related protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hashimoto M., Ichihara M., Takahashi M.;
RT   "Cloning and characterization of GPI-anchored alpha-2 macroglobulin-
RT   related protein.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY083458; AAM00021.1; -.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
SQ   SEQUENCE   1442 AA;  161658 MW;  E2B9671B8EA5BAA7 CRC64;

  Query Match             72.9%;  Score 5360;  DB 11;  Length 1442;
  Best Local Similarity   71.6%;  Pred. No. 4.1e-318;
  Matches 1038;  Conservative  176;  Mismatches  202;  Indels   34;  Gaps    8;

Qy         1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
             |:   ||:||||||:|  ||| ||| ||||||||||||| ||||||:|||: | || |||
Db         1 MRSRRLLSAAHLLCLCAVALA-APGSRFLVTAPGIIRPGANVTIGVDLLENSPPQVLVKA 59

Qy        61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
             :: | ||| : |:||||||| :| |||| ||:|||:|||:|||| : |::::||:||| |
Db        60 QVFKIASNKSRSILEAEGVFHRGHFKTLVLPALPLSSADKIYELHINGQSENEIVFSNRT 119

Qy       121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
             ||:||:| ||| |||||| ||||||||||::|| || |||:||::| |||||||:||||
Db       120 RLTFESKSISVLIQTDKAFYKPKQEVKFRVLTLCSDLKPYRTSVDIFIKDPKSNVIQQWF 179

Qy       181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
             ||: ||||:||||||||:|| |||||||||||| ||||||| ||||||||||:|||||||
Db       180 SQKGDLGVVSKTFQLSSNPIFGDWSIQVQVNDQQYYQSFQVLEYVLPKFEVTVQTPLYCS 239

Qy       241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
             : || |||:: ||||||||||| ::|||||||||||||||||:|:||| |||||:: |||
Db       240 LKSKQLNGSVIAKYTYGKPVKGSLSLTFLPLSFWGKKKNITKSFEINGFANFSFDNYEMK 299

Qy       301 NVM------DSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFF 354
```

```
                 ||       | | |  | :| | ||| ||: ||||:|||||  ||||||||||||| |
Db        300 KVMNLKPLTDVSEGSYENVDPSFPGPAEIIATVTESLTGISRMASTNVFFKQHDYIIEIF 359

Qy        355 DYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQR--NYTEYWSGSNSGNQKM 412
              ||||||||||||||||||:|:||||| ||  |::|  |||  |: |          : :|:|
Db        360 DYTTVLKPSLNFTATVKVSRSDGNQLTPEEIENDLVTVVTQRKNNHPE-----SQRDQEM 414

Qy        413 EAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKT 472
              : :| :||:||:|  |||||::   |||||||||  ||: |||:  ||||||||||
Db        415 DYIQTVNYTIPQNGIIKIEFPVMSISGELQLKAYFLDGTSSVTVHSMFTSPSKTYIQLKT 474

Qy        473 RDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVI 532
              ||| ||||||:|:|||:: |:|||||:|:||||| |||:|   |||||| || ||||:|
Db        475 RDEYIKVGSPFDLMVSGNRQFKDLSYMVISKGQLVAAGKQSSRTFSLTPEASWAPKACII 534

Qy        533 VYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAV 592
               ||| :||||:|:|||||||||:|:|:||:|||     :||:||||||| || ||:||||||
Db        535 AYYIAEDGEIINDILKIPVQLVFENKVKLFWSKPTVKPSDKVSLRISATQSDSLVGIVAV 594

Qy        593 DKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYID 652
              |||| ||  || |||| :||||||||| |||||||||||||||||||||||| | :| ||
Db        595 DKSVTLMENSNSITMETMVHELELYNTEYYLGMFMNSFAVFQECGLWVLTDATLIRDSID 654

Qy        653 GVYDNAEYAERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVT 712
               ||  ||:||| ||||  ||||  ::||    | |     ::  |||:||||||  |  :|:||||
Db        655 EVYDTEEYSERFAEENEANLVDFEDASSVNNVHVRKNFPETWIWLDAYMGSKIYEEFEVT 714

Qy        713 VPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFN 772
              |||||||||:  |||||||  ||||  |  ||||||||||:|||||||||||||||||::| |
Db        715 VPDSITSWVASAFVISEDLGFGLTTVPAELQAFQPFFLFLNLPYSVIRGEEFALEVSIVN 774

Qy        773 YLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPI 832
              ||||  :| ::||:||  ||||||::  |  |  :::|: || ::| |::|||:||||||||
Db        775 YLKDTIKVVILIEESDSFDILMTSNDTNGTIYRKTVQVPRDNGVTLVFPIKPTHLGEIPI 834

Qy        833 TVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGS 892
              |||| |||||||||  |:||  ||||||||:|:|||||:  ::|    :::  |||||:|| ||
Db        835 TVTAASPTASDAVTQTIVVKPEGIEKSYSKSVLLDLTDSNVESKQQSMRFSFPPDTVIGS 894

Qy        893 ERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEK 952
              |:||||||||:|| ||||| |||||||||||||||  |||||||||||| |:|||| |||||
Db        895 ERVQITAIGDILGSSINGLSSLIRMPYGCGEQNMIYFAPNIYILDYLTKQKQLTVNLKEK 954

Qy        953 ALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHR 1012
              |||:||||||||||||||||||||||  ||||||||||||||||||||| |||||||:||||
Db        955 ALSYMRQGYQRELLYQREDGSFSAFGDIDSSGSTWLSAFVLRCFLEADYYIDIDQDVLHR 1014

Qy       1013 TYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESI 1072
              |||||  |:| ||||:||||||||||| |||||||||||||:|||:||||||||||:|
Db       1015 TYTWLNAHKKFNGEFWEPGRVIHSELQGGTKSPVTLTAYIVTSVLGYKKYQPNIDVQDSI 1074

Qy       1073 HFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLS 1132
               |||  |||||||||||||:|:||||:||||||:||||:|  |:|:||  |||:||  ||
Db       1075 KFLEFEFSRGISDNYTLAIISYALSTVGSPKAEEALNLLMQRSEKEGDTQFWLSSGPALS 1134

Qy       1133 DSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALS 1192
               ||||||:|||:||||||:| |   ||||:||| :|||||||| |||| ||||||
Db       1135 GSWQPRSVDIEIAAYALLAHTLH-HVSEGIPVMRWLIQQRNSLGGFVSTQDTVVALKALS 1193

Qy       1193 EFAALMNTERTNIQVTVTGPSSP-----------------SPLAVVQPMAVNISANGFGF 1235
              ||:||::   |:|||:|||||                   :  |:  |:||:|||  |
Db       1194 EFSALVHKENTDIQLTVTGPGIPRSIHFRIDSQNLFLLHQEELHALDPITVNVSAHGSGF 1253

Qy       1236 AICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSG 1295
              |||||| |||| ||||:||||  :||  ||||| |   |:||::|::||||||  |  |:|
Db       1254 AICQLNVDYNVKGSGSSKRRRSTENQEVFDLDVIV-NNEDDISHLNLNVCTSHLGSERTG 1312

Qy       1296 MALMEVNLLSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVS 1355
               | |||||||||      |::|  ||||:||||||:||||||||||||:|||||:||  ||::|||
Db       1313 MVLMEVNLLSGFSASSDSIPLSETLKKVEYDNGKLNLYLDSVNESQFCVNIPTVRDYKVS 1372

Qy       1356 NTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFI 1415
               | :| | |||::|||||||||||||||::||||||  |   |    |:    |||:  |   || : :
Db       1373 NIRDGSVSVMDYYEPRRQAVRSYNTQVKLSSCYLSPDT-NCKSHTDGATDSLRRSSSLLV 1431

Qy       1416 FCFKLLYFME 1425
              ||  ||||::
Db       1432 FCSVLLYFVQ 1441
```

RESULT 3
Q8N3A7

```
ID   Q8N3A7      PRELIMINARY;      PRT;   854 AA.
AC   Q8N3A7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   DKFZP762L1111.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Melanoma;
RA   Blum H., Bauersachs S., Mewes H.W., Weil B., Wiemann S.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL834478; CAD39137.1; -.
DR   InterPro; IPR001599; Macroglob1nA2.
DR   Pfam; PF00207; A2M; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   854 AA;  95599 MW;  7B51863A2565C004 CRC64;

  Query Match             58.6%;  Score 4305.5;  DB 4;  Length 854;
  Best Local Similarity   97.7%;  Pred. No. 4.7e-254;
  Matches  834;  Conservative    1;  Mismatches    2;  Indels   17;  Gaps    1;

Qy        592 VDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYI 651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYI 60

Qy        652 DGVYDNAEYAERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEV 711
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db         61 DGVYDNAEYAERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGSRIYQEFEV 120

Qy        712 TVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIF 771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIF 180

Qy        772 NYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIP 831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 NYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIP 240

Qy        832 ITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTG 891
              ||||||||||||:|||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ITVTALSPTASDAITQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTG 300

Qy        892 SERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKE 951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 SERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKE 360

Qy        952 KALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLH 1011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLH 420

Qy       1012 RTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQES 1071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 RTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQES 480

Qy       1072 IHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKL 1131
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 IHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKL 540
```

```
Qy       1132 SDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKAL 1191
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 SDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKAL 600

Qy       1192 SEFAALMNTERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFG 1234
              ||||||||||||||||||||||||||                 |||||  |||||||||
Db        601 SEFAALMNTERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPTAVNISANGFG 660

Qy       1235 FAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRS 1294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 FAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRS 720

Qy       1295 GMALMEVNLLSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKV 1354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GMALMEVNLLSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKV 780

Qy       1355 SNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIF 1414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIF 840

Qy       1415 IFCFKLLYFMELWL 1428
              ||||||||||||||
Db        841 IFCFKLLYFMELWL 854


RESULT 4
Q8N915
ID   Q8N915     PRELIMINARY;     PRT;   665 AA.
AC   Q8N915;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein FLJ38569.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA   Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA   Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA   Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA   Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA   Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA   Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK095888; BAC04642.1; -.
DR   InterPro; IPR002890; A2M_N.
DR   Pfam; PF01835; A2M_N; 2.
KW   Hypothetical protein.
SQ   SEQUENCE   665 AA;  74411 MW;  CC73982C8E88647E CRC64;

  Query Match             45.7%;  Score 3360;  DB 4;  Length 665;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-196;
  Matches  657;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDN 657
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDN 657


RESULT 5
O46015
ID   O46015     PRELIMINARY;     PRT;  1519 AA.
AC   O46015; O46014;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical 168.4 kDa protein ZK337.1.
GN   ZK337.1.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   White S.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ZK337.1B (SHOWN HERE) AND
CC       ZK337.1A; ARE PRODUCED BY ALTERNATIVE SPLICING.
DR   EMBL; Z82090; CAB05007.1; -.
DR   EMBL; Z82090; CAB05006.1; -.
DR   HSSP; P01024; 1C3D.
DR   WormPep; ZK337.1a; CE16718.
DR   WormPep; ZK337.1b; CE16719.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR000847; HTH_LysR.
DR   InterPro; IPR001599; Macroglob1nA2.
DR   InterPro; IPR001917; NHtransf_2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW   Alternative splicing; Hypothetical protein.
FT   VARSPLIC    714    724       MISSING (IN ISOFORM ZK337.1A).
SQ   SEQUENCE   1519 AA;  168385 MW;  78139ACB564AB742 CRC64;

  Query Match             25.2%;  Score 1850;  DB 5;  Length 1519;
  Best Local Similarity   31.6%;  Pred. No. 8.6e-104;
  Matches  477;  Conservative  290;  Mismatches  563;  Indels  180;  Gaps   44;

Qy         15 VCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVL 74
              | | |  | |   ::: || ::||    :: : ||:    :   :     ::|  | |::
```

```
Db        28 VSTTAAPVKPA-TYMLVAPAVVRPDQPFSVCMNLLKQATDEDMIVRIEVRTERNETIAAR 86

Qy        75 EAEGVFEKGSFKTLTLPSLPLNSAD--EIYELRVTGRTQD-EILFSNSTRLSFETKRISV 131

Db        87 VISNL-KPGIAQTVSLSEMPAQSLTPRQSYKLYIRGETLNAELIFENENELKYDQKALSV 145

Qy       132 FIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISK 191

Db       146 FIQTDRAIYRPASLVRYRAIVVKSDLKPYVGNATIKIFDPSRNLISQTIGVTLDRGVYSG 205

Qy       192 TFQLSSHPILGDWSIQVQ----VNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLN 247

Db       206 ELQLAEETLLGDWFIEVETSNGVQDKS---SFTVDTYVLPKFEVNIKTSSFITIND-DLS 261

Qy       248 GTITAKYTYGKPVKGDVTLTF-LPLSFW-----------GKKKN----ITKTFKIN--GS 289

Db       262 VFVDAKYTYGKGVAGKAKVSLELPWHRWHAMVPTIIDENGVKKEEELMVERTVKLNRQGE 321

Qy       290 ANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDY 349

Db       322 AAVVFSNDELKR--------HKLLHEWGGGSIRIVASVTEDITEIERNATHQISTFREEV 373

Qy       350 IIEFFDYTTVLKPSLNFTATVKVTRADGN--QLTLEERRNNVVITVTQRNYTEYWSGSNS 407

Db       374 KLDVEKQGDTFKPGLTYNVVVALKQMDDTPVKATLPKR----VQVSTFYNYP--YNHDTS 427

Qy       408 GNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK------ 461

Db       428 SLQEEKETKIVEVDAHGTSVLTLQPPINCTSARIEAH-YDIGGKDNFTATPIYSSLYVEA 486

Qy       462 --SPSKTYIQLKTRDEN-IKVGSPFELVVSGNKRLKELSYMVVSRGQLV-----AVGKQN 513

Db       487 AVSPTKSFLQLLADNEGAVDVGKSLSFSLKATQPLSTITYQVMSRSNIVVSQQMTVNSEH 546

Qy       514 STMFSLTPENSWTPKACVIVY-YIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSE 572

Db       547 ATI-SFPATANMAPKSRLIVYAIIESSQEVLVDALDFKVEGIFQNQVALSIDKQAVEPGQ 605

Qy       573 KVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMF------ 626

Db       606 NVKFKVT-SDKNSFVGLLVVDQSVLLLKTGNDITREKVEQDLENYDSNNVGGGFGGPRPW 664

Qy       627 -------------------MNSFAVFQECGLWVLTDA--------------------NLT 647

Db       665 EAIDRKKRSIWRPWWGIGGSDAQSIFSNAGLVVLTDALLYREPQREFMSERRLNTPGGLT 724

Qy       648 KDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGSS---PHVRKHFPETWIWLDTNMGYR 704

Db       725 VMMMDGAPGMAEAA--FAAPPMG----------GSSPPPPTVRKFFPHTWIWSDLN-STS 771

Qy       705 IYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEF 764

Db       772 GEVEMEIEAPDTITSWVASTFAINEENGLGVAPTTSKLRVFRPFFIQLNLPYAVRRGEKF 831

Qy       765 ALEITIFNYLKDATEVKVII--EKSDKFDILMTSSEI---NATGHQQTLLVPSEDGAT-- 817

Db       832 ALLVLVFNYMEKEQDVTVTLKYDKDSGYDLLKKDGTVVRRDEVGQQNVRIVSVAGGGTSK 891

Qy       818 -VLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQST 876

Db       892 AVYFPIVPSSIGEIPVHISAIASQGGDAVEMNLRVDPQGYKVDRNIPFVIDLNNNSSDFS 951

Qy       877 LKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYIL 936

Db       952 -KNLELIWPNDVVDGSQKARLDVIGDMMGPVLNNAHKLVQMPYGCGEQNMLNLVPNILVV 1010

Qy       937 DYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF 996

Db      1011 KYLRATNRNESQLETKAIKFIEQGIQRELTYKRADNSFSAFGDSDKAGSTWLTAFVVRSF 1070

Qy       997 LEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG-NKSPVTLTAYIVTS 1055

Db      1071 HHAKQYAFVDPNVISRAVAFLNSQQMESGAFAERGEVHHKDMQGGAQDGGVALTAFVLIS 1130

Qy      1056 LLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWR- 1114

Db      1131 IL-----ENGMENGKAVTYLEKHLDEVSGNAYTMAVVAYALQLAKSKQAGKAFENLKKHK 1185

Qy      1115 -AEQEGGMQFWVSSE--SKLSDS----WQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRW 1167

Db      1186 IVEKSGDVKFASAQKKVEKLKESRAYMFQARPVDIETTSYAVLSYLAQNQTSESLSIIRW 1245

Qy      1168 LSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTV-TGPSSPS------PLAV 1220

Db      1246 LVSQRNELGGFTSTQDTVMALQALSSYAAVTYSDKHTSQVTILNGKHTHSFDINIRNAIV 1305

Qy      1221 VQPM-------AVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKEN 1273

Db      1306 LQSYQLSSLNDAVSINANGTGVVFAQLSYSY---------YRDSLNDDAPFFCSQEIKEI 1356

Qy      1274 KDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVK-----KVEYDHG 1328

Db      1357 RAG-NRLQLDLCCNYTRPGKSNMALAEIDALSGYRFDAEQVHTLTSIEDLQRVEMEKDDT 1415

Qy      1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCD 1388

Db      1416 KMNVYFNPLGGRPVCLSLYSDVTYQVADQKPANFRLVDYYDPEEQLKMTYAAKQTRSLQE 1475

Qy      1389 LCSDVQGCRP 1398

Db      1476 KCG--EDCWP 1483
```

RESULT 6

```
Q8T398
ID   Q8T398      PRELIMINARY;      PRT;  1809 AA.
AC   Q8T398;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Alpha-2-macroglobulin homologue.
GN   A2M.
OS   Ciona intestinalis.
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC   Phlebobranchia; Cionidae; Ciona.
OX   NCBI_TaxID=7719;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Hepatopancreas;
RA   Hammond J.A., Nakao M., Yano T., Kemp G.D., Smith V.J.;
RT   "Complete cDNA of an alpha2 macroglobulin homologue with a catalytic
RT   histidine residue from Ciona intestinalis.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ431688; CAD24311.1; -.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR001599; Macroglobln A2.
DR   InterPro; IPR000884; TSP1.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   SMART; SM00209; TSP1; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR   PROSITE; PS50092; TSP1; 1.
SQ   SEQUENCE   1809 AA;  199172 MW;  E300DC643946E454 CRC64;

  Query Match             24.2%;  Score 1779.5;  DB 5;  Length 1809;
  Best Local Similarity   30.0%;  Pred. No. 2.3e-99;
  Matches  453;  Conservative  293;  Mismatches  586;  Indels  179;  Gaps   41;

Qy        13 LCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVS 72

Db        16 LLATLSSLATASNV-YNIYFPKHIRPGFNISFTAAIIDN-PNTVQIHTAFRSMDNSFHVD 73

Qy        73 VLEAEGVFEKGSFKTLTLPSLPLN-SADEIYELRVTGR---TQDEILFSNSTRLSFETKR 128
```

```
Db         74 STDS---VNSGSSSRISMNGLPIHYSGSHGFELNITGTDLVTGAQLFFNSSTDFQFQAKS 130

Qy        129 ISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGV 188
              ||: |||||:|:|     |||| : |   | || : :::     |||: |::         : ||
Db        131 ISILIQTDKAIYQPGHTVKFRAIALKPDLKPLQGNISYTFKDPRGNVVMLEPEVPLNHGV 190

Qy        189 ISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLNG 248
                | |:   :  | | ::              | :|   | ||||:| :: | |    |   |
Db        191 AGGQFSLTKDAVAGMWKVEFMAEGFKESLSVEVKRYKLPKFKVEVKAPSYIHPQSTGLTI 250

Qy        249 TITAKYTYGKPVKGDVTLTFL-----PL--SFWGK-------KKNITKTF-KINGSANFS 293
              : ||||:|| |:|     |   :             |:     | |:           :  ||:  :   :|:
Db        251 KLDAKYTFGKGVQGTGLLEVVGGYQYPVYHGFGGRFAPRPPTQNKITRRYPNFDGTVELL 310

Qy        294 FNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEF 353
                ::|::   :     || || :            : :   :|||::|   :  |  :    :   |    :   :|
Db        311 ITNDEIREEL-GWNGASESI-------ITVTGSVTEALTREAFNDTQRIDAKTTNVKVET 362

Qy        354 FDYTTVLKPSLNFTATVKVTRADGNQLTLEER-RNNVVITVTQRN--------------- 397
                   :|| |  ::|  :::|   ||   |   ::|  ||::: :   |
Db        363 LVKPLTIKPGLKYSAYIQITEVDGKPLPEDDRLANNLLLNIEYRYPRGEPEPGTNTTVST 422

Qy        398 -YTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDS-SELQLKAYFLGSKSSMA 455
                |    |    :                      : :|  ||    |:           |: : :   :   |     : ::|:
Db        423 WYAYRWEETRV------------FVIPPSGIVKVTIDAPSDTFTSINFRPY---TNATMS 467

Qy        456 VHSLFK-------SPSKTYIQLKTRDENIKVGSPFFELVVSGNKRLKELSYMVVSRGQLVA 508
                   :             ||| :|:|:|  | : :      :       :   :      :   | : :::|||::::
Db        468 QRWALQWTAERADSPSNSYLQITTEENSVVPGNMATVTIRTTEAVSEFTILIISRGEILS 527

Qy        509 VGK--------QNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIK 560
                  |             :|| :|   :  |     |     |:    |:  |||||::| :|:  |        :|::
Db        528 ERKFQTLSGVPENSHLFEFSVEYDMIPGVQVLASYVRDDGEIVADYIKLTVTAELENQVS 587

Qy        561 LYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTG 620
               :    |      :     |||:|:    :      :  ||     |:|:||    |: : |   ||::  |   :|  :|     |:
Db        588 ITSSSTNIDAGEDVSIRVQTSSSGAYVGARAIDQSVLLLKSGNDVSQERIVTDLNKYSVT 647

Qy        621 YYLGMF-------------MNSFAVFQECGLWVLTDANLTKDYIDGVYD--------NAE 659
                 |                      ::  ||:: |: | |||  : :         :|           |
Db        648 QELNHMWRWWWWYPTPSGASDASDVFRKAGILVFTDALVYQKPEASIYPFRPIAFSLNGG 707

Qy        660 YAERFMEENEGHIVDIH--DFSLGSSP-HVRKHFPETWIWLDTNMGYRIYQEFEVTVPDS 716
               :|||               :|:            | |      ::|        |||||:|  :            |       |  ||:
Db        708 FAER-------NIIATAAVDTSTPATPTRTRTLFPETWLWDEQISGADGSATFNTTAPDT 760

Qy        717 ITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKD 776
              ||||: : | :|   |||::    ::     |:   |||  ||||  |||||    ::  :||||
Db        761 ITSWIFSAFSVSDQHGLGVSEQH-KVTVFRNFFITLNLPVRVIRGELIIVQAIVFNYLST 819

Qy        777 ATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTA 836
                 :  :  : : :|:|| :|      :       |  |   : : :|:    :|  ||||     |||||||:||
Db        820 EVDAVLTLTESNKFVLLRPGNNSAAVGFSRRITIPASGSVSVKFPIRMGTLGEIPITMTA 879

Qy        837 LSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
               :|  ||||:|: :  |:  |||  :   |  |:|      |          :::|:     |     | |||:|:
Db        880 ISEIASDALTRKVFVQPEGITQCTSGSVLFQRMDASAPPDVESLNIQIPAGIVPGSEKVK 939

Qy        897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
               :     ||:|| ::| | ||:| | |||||||: |||::::  ||          :|           :  ||
Db        940 LLVYGDILGSTMNNLGSLLRTPSGCGEQNMLGFAPDVFVTLYLHSAGKLDAATRAKAFKH 999

Qy        957 MRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFL---EADPYIDIDQNVLHRT 1013
               :  || ||  |  |:   |||||||| | |||||||:||   |   | :  |           :  : |   :  :
Db       1000 FQTGYSNELNYKHRDGSFSAFGEGDASGSTWLTAFAAKCFMFARELRPTL-VSASVIDQA 1058

Qy       1014 YTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIH 1073
               |:|    |  :  |  |  :||||  |     :|||   ||:|:|||::  :|               |     |||:
Db       1059 LTFLINQQNTTGTFREPGRVSHKAMQGGVDSPITMTAYVLITLKETNYAVKNRAVQEAAE 1118

Qy       1074 ----FLESEFSRGISDN-YTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFW---V 1125
                   :||:    :   |||| | ||::||||   || :|  |||    |     |    :||  :||
Db       1119 NARIYLENHLT-SISDNKYALAIVTYALHVAGSSRANEALLALEALATVQGGFKFWHDNS 1177

Qy       1126 SSESKLSDSWQ-----PRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFAS 1180
                  |     |   |:        | : |||::||||||::   :     :   |||:|:||: :|:||||::
Db       1178 ESPDSYSSRWRPYYYNPPTNDIEMSAYALLTYVRRNDLNAGIPVMKWLASKRSSLGGYSG 1237

Qy       1181 TQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPL------------AVVQPM--- 1224
                  |||| :|::|||: |  |:       |:|:: :   :   |                     :|   |
Db       1238 TQDTVIAIQALSKVAGLLVGNTQNLQISASHSNDPFTASYNINRENSIVFNSVNVPAVDG 1297

Qy       1225 AVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQ----EAFD-----LDVAVKENKD 1275
               | ::|  |  |  |  |:  |::|  ||                      ||       |  |        :     |:|: |
Db       1298 TVQVTATGVGVAVAQISVCYNT-----------PNQPYEIEPFQCTNTVVSTALKKAK- 1344

Qy       1276 DLNHVDLNVCTSF-SGPGRSGMALMEVNLLSGFMVP-SEAISLSETVKKVEYDHGKLNLY 1333
                  :|  |   |    |      :||  |||||  ||: |          :  :  : |  ||  |      :|:|
Db       1345 ------VNWCCSLRPGDNATGMFLMEVNLPSGYTVNIDNERTRNPSAKLVEIDGNGVNVY 1398

Qy       1334 LDSVNE-TQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCS- 1391
                  | :      | :|   :      |   ::      |:   |||:|: :      |      :         ||  ||
Db       1399 YDELAPGRSVCADIELLNLGNVGGSKARKVAASDYYQPKERVEALYQVDEAPVVCDSCST 1458

Qy       1392 -DVQGCRPCED 1401
                 |:  |   |  |
Db       1459 EDIAVCSVCAD 1469
```

RESULT 7
Q8IPH5

```
ID   Q8IPH5      PRELIMINARY;      PRT;  1388 AA.
AC   Q8IPH5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG7052-PE.
GN   TEPII.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
```

```
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003618; AAN10639.1; -.
SQ   SEQUENCE   1388 AA;  154276 MW;  FE257D37D7DDCE1E CRC64;

  Query Match             23.2%;  Score 1703;  DB 5;  Length 1388;
  Best Local Similarity   31.2%;  Pred. No. 7e-95;
  Matches  458;  Conservative  288;  Mismatches  539;  Indels  182;  Gaps   50;

Qy         19 ALAVAPGPRFLVTAPGIIRPGG--NVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEA 76
Db         14 ALLVNATGIYSVVGPGTLRSNSKYNVVVSVHKADG-PSQ--IKVSLNGPSYNET------ 64

Qy         77 EGVFEKGSFKTLTLPSLPLNSADEIYEL--------RVTGRTQDEILFSNSTRLSFETKR 128
Db         65 ---------KQIELP--PMSTQNVEFEVPKLATGNYNLSAEGVSGVVFKNSTKLNYADKK 113

Qy        129 ISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTS--LNILIKDPKSNLIQQWLSQQSDL 186
Db        114 PSVFVQTDKATYKPADLVQFRILFLDENTRPAKIEKPISVIIIDGAQNRIKQLSDVKLTK 173

Qy        187 GVISKTFQLSSHPILGDWSIQVQVN-DQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKH 245
Db        174 GVFSGELQLSEQPVLGTWKISVSVDGDNRETKSFEVDKYVLPKFEVIVDTPKAVVIADKV 233

Qy        246 LNGTITAKYTYGKPVKGDVTLT-------FLPLSFWGKKKNITKTFKINGSANFSFNDEE 298
Db        234 IKATIRAKYTYGKPVKGKATVSMERSYGYFGDLNANGNKQE--KTIDVDGKGHVEFD--- 288

Qy        299 MKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTT 358
Db        289 --IIHWAQRG--QYLP-----PIKLFAVVTEELTGNKQNATATVVLHQQRYSIEPYERPE 339

Qy        359 VLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKI 418
Db        340 HFEANKSFIYQVVVKNVDGSPVT--NSAKNVKIGF-DKSYSYFHEPSPK--------TRI 388

Qy        419 NYTVPQS----GTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK-SP-SKTYIQLKT 472
Db        389 NFEAPVNENGIATFNVRLPD-SDSRYYRIFASFDGSENTIGSISKFEPTPMSREPLKIQV 447

Qy        473 RDENIKVGSPFELVVSGNKRLKELSYMVVSRGQL-------VAVGKQNSTMFSLTPENSW 525
Db        448 NTKKPRLGEQVSFDVVSIEDLPYFVYTIVARGNVILSDYVDVPDGQKTYTV-KFTPTFSM 506

Qy        526 TPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDS 585
Db        507 VPKATIYVYYVVNN-DLQFEEKTIDFEKEFSNSIDVS-APTNAKPSEEVKLRIK-TDADS 563

Qy        586 IVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVF--QECGLWVLTD 643
Db        564 FVGLLGVDQSVLLLKSGNDLSQDDIFNSLNIYQTS---TPWMNGYGRYPGQTSGLVTLTN 620

Qy        644 ANLTKDYIDGV-YDNAEYAERFMEENEGHIVDIHDFSLGSSPH---VRKHFPETWIWLDT 699
Db        621 ANYPYNTEDSISYD----------------EVDAISITSSTKIELVRTNFAEVWMWTTS 663

Qy        700 NMGYRIYQEFEVT--VPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYS 757
Db        664 DNGSVGEEEFTLTKKIPDTITSWVVTGFSLNPTSGIALTKNPSKIRVFQPFFVSTNLPYS 723

Qy        758 VIRGEEFALEITIFNYLKDATEVKVIIEKSD-KFDILMTSSE-----INATGHQQTLLVP 811
Db        724 VKRGEVIAIPVVIFNYLDKTLDADVVMDNSDQEYEFTEATNEVLEKAIDEVRRVKRVTIP 783

Qy        812 SEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDN 871
Db        784 ANSGKSVSFMIRPKNVGFTTLKITATSALAGDAIHQKLKVEPEGVTLFENRAVFINLKDQ 843

Qy        872 RLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAP 931
Db        844 PEMS--QSLDADIPNEVVPQSEFIEFSVVGDLLGPTLQNLDNLVRMPYGCGEQNMVNFVP 901

Qy        932 NIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAF 991
Db        902 NILVLKYLEVTGRKLPSVESKARKFLEIGYQRELTYKHDDGSYSAFGKSDASGSTWLTAY 961

Qy        992 VLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAY 1051
Db        962 VMRSFHQAGTYTDIDPKVITAGLDFLVSKQKESGEFPEVGKL----FDNANQNPLALTSF 1017

Qy       1052 IVTSLLGYRKYQPNID--VQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALN 1109
Db       1018 VLLAFFENHELIPKYQSAIKKAVRYVAEEADK-TDDQYSLAIAAVALQLAKHPQSEKVIA 1076

Qy       1110 MLTWRAEQEGGMQFWVSSESKLSDS---------WQPRSLDIEVAAYALLSHFLQFQTSE 1160
Db       1077 KLESVARKENDRMWW----SKATESTGEDGRVFHWKPRSNDVEITSYVLLALLEKDPAEK 1132

Qy       1161 GIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFA--ALMNTERTNIQVTVTGPS----- 1213
Db       1133 ALPIIKWLISQRNSNGGFSSTQDTVIGLQALTKFAYKTGSGSGTMDIEFSSAGESKNTIK 1192

Qy       1214 -SPSPLAVVQ-------PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFD 1265
Db       1193 VNPENSLVLQTHDLPKSTRKVDFTAKGTGSAMVQLSYRYNLAEK---------EKKPSFK 1243

Qy       1266 LDVAVKENKDDLNHVDLNVCTSF------SGPGRSGMALMEVNLLSGFMVPSEA---ISL 1316
Db       1244 VTPTVKDTPNQLLIVD--VCAEYVPLEDADKDKDSNMAVMEIALPSGFVGDSTSLGKIQA 1301

Qy       1317 SETVKKVEYDH--GKLNLYLDSVNETQF-CVNIPAVRNFKVSNTQDASVSIVDYYEPRRQ 1373
Db       1302 VDRVKRVETKNSDSTVVVYFDSLTPGDVRCLPLEASKAHAVAKQKPASVSLYDYYDTERK 1361
```

```
Qy       1374 AVRSYNSEVKLSSCDLCSDV---QGCR 1397
              |.  |  :|| | ||:|      :||:
Db       1362 ATEYY--QVKSSLCDICEGADCGEGCK 1386


RESULT 8
Q9VLZ0
ID   Q9VLZ0      PRELIMINARY;      PRT;  1397 AA.
AC   Q9VLZ0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG7052 protein.
GN   TEPII OR CG7052 OR CG18589.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003618; AAF52540.2; -.
DR   HSSP; P01024; 1C3D.
DR   FlyBase; FBgn0041182; TepII.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   InterPro; IPR000215; Serpin.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS00284; SERPIN; 1.
SQ   SEQUENCE   1397 AA;  155108 MW;  9CE9B684AB207574 CRC64;

  Query Match             23.1%;  Score 1699.5;  DB 5;  Length 1397;
  Best Local Similarity   31.3%;  Pred. No. 1.2e-94;
  Matches  460;  Conservative  288;  Mismatches  543;  Indels  179;  Gaps   52;

Qy         19 ALAVAPGPRFLVTAPGIIRPGG--NVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEA 76
              || |       : |  || :|      || : |    :  |||  :|  |    : | |
Db         14 ALLVNATGIYSVVGPGTLRSNSKYNVVVSVHKADG-PSQ--IKVSLNGPSYNET------ 64

Qy         77 EGVFEKGSFKTLTLPSLPLNSADEIYEL--------RVTGRTQDEILFSNSTRLSFETKR 128
                        | :  ||  |::: :  :|:              ::   ::| ||:|:: |:
Db         65 ---------KQIELP--PMSTQNVEFEVPKLATGNYNLSAEGVSGVVFKNSTKLNYADKK 113

Qy        129 ISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTS--LNILIKDPKSNLIQQWLSQQSDL 186
              |||:||||| |||    |:|||: |  :  :|  |     ::::|  |   | |:|
Db        114 PSVFVQTDKATYKPADLVQFRILFLDENTRPAKIEKPISVIIIDGAQNRIKQLSDVKLTK 173

Qy        187 GVISKTFQLSSHPILGDWSIQVQVN-DQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKH 245
              || |    ||  |:||  | | | |: |    :||:| :|||||||| : ||     :  |
Db        174 GVFSGELQLSEQPVLGTWKISVSVDGDNRETKSFEVDKYVLPKFEVIVDTPKAVVIADKV 233

Qy        246 LNGTITAKYTYGKPVKGDVTLT-------FLPLSFWGKKKNITKTFKINGSANFSFNDEE 298
              :  || ||||||||||  |::       |  |:  | |:   ||   ::|   :  |:
Db        234 IKATIRAKYTYGKPVKGKATVSMERSYGYFGDLNANGNKQE--KTIDVDGKGHVEFD--- 288

Qy        299 MKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTT 358
                  :  :   |   :||      |:::     ||| :||. :| :     |   | || ::
Db        289 --IIHWAQRG--QYLP-----PIKLFAVVTEELTGNKQNATATVVLHQQRYSIEPYERPE 339

Qy        359 VLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKI 418
              ·: : :|   |  |    ||: :|          ||  |      ::|: :    |      :|
Db        340 HFEANKSFIYQVVVKNVDGSPVT--NSAKNVKIGF-DKSYSYFHEPSPK--------TRI 388

Qy        419 NYTVPQS----GTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK-SP-SKTYIQLKT 472
              |:  | :        ||  :    ||     :: | | ||::::     | :  :| |:  ::::
Db        389 NFEAPVNENGIATFNVRLPD-SDSRYYRIFASFDGSENTIGSISKFEPTPMSREPLKIQV 447
```

```
Qy        473 RDENIKVGSPFELVVSGNKRLKELSYMVVSRGQL-------VAVGKQNSTMFSLTPENSW 525
                :  ::|        |    : |      | :|:|| :          |  |::  |:   || |
Db        448 NTKKPRLGEQVSFDVVSIEDLPYFVYTIVARGNVILSDYVDVPDGQKTYTV-KFTPTFSM 506

Qy        526 TPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDS 585
              ||| : |||: :: ::   :    ·|  :   | | | :   :     |:|||:| |||  |  ||
Db        507 VPKATIYVYYVVNN-DLQFEEKTIDFEKEFSNSIDVS-APTNAKPSEEVKLRIK-TDADS 563

Qy        586 IVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVF--QECGLWVLTD 643
               ||:: ||:|| |: : ||:: ::: : | :| |         :|| :  :  |  || ||:
Db        564 FVGLLGVDQSVLLLKSGNDLSQDDIFNSLNIYQTS---TPWMNGYGRYPGQTSGLVTLTN 620

Qy        644 ANLTKDYIDGVYDNAEYAERF-MEENEGHIVDIHDFSL-----GSSPHVRKHFPETWIWL 697
              ||              | |:   |  :  |  :   |           :| :|| ||| ||:
Db        621 ANYPY--------NTEFPIAFSLAAPQAAIAGMPGTSSIASHPNQAPQIRKEFPENWIFY 672

Qy        698 DT-NMGYRIYQEFEVT--VPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNL 754
              :   |:|    :|| :|   :||:||||| ||| ::      |: ||   | ::: |||||:  ||
Db        673 NAENVG---EEEFTLTKKIPDTITSWVVTGFSLNPTSGIALTKNPSKIRVFQPFFVSTNL 729

Qy        755 PYSVIRGEEFALEITIFNYLKDATEVKVIIEKSD-KFDILMTSSE-----INATGHQQTL 808
              |||| |||  |: : ||||      | :::  || :::     ::|        |:      :  :
Db        730 PYSVKRGEVIAIPVVIFNYLDKTLDADVVMDNSDQEYEFTEATNEVLEKAIDEVRRVKRV 789

Qy        809 LVPSEDGATVLFPIRPTHLGEIPITVTLALSPTASDAVTQMILVKAEGIEKSYSQSILLDL 868
               :|:   | :|  | ||| ::|     : :|| |     | ||: | : |: ||:       :::: ::|
Db        790 TIPANSGKSVSFMIRPKNVGFTTLKITATSALAGDAIHQKLKVEPEGVTLFENRAVFINL 849

Qy        869 TDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMIN 928
               |     |  ::|    |    |    || :: : :||:|||::   | :|:||||||||||:|
Db        850 KDQPEMS--QSLDADIPNEVVPQSEFIEFSVVGDLLGPTLQNLDNLVRMPYGCGEQNMVN 907

Qy        929 FAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWL 988
              | ||| :| ||      :     ::: ||   |:  ||||| |: :|||:||||   | ||||||
Db        908 FVPNILVLKYLEVTGRKLPSVESKARKFLEIGYQRELTYKHDDGSYSAFGKSDASGSTWL 967

Qy        989 SAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTL 1048
              :|:|:| | :|  | |||  |:      | ||  :  | ||| : |::              |::|: |
Db        968 TAYVMRSFHQAGTYTDIDPKVITAGLDFLVSKQKESGEFPEVGKL----FDNANQNPLAL 1023

Qy       1049 TAYIVTSLLGYRKYQPNID--VQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKE 1106
              |:::: :      :     |    ::::: ::   :   |  | |:||:  ||    |::::
Db       1024 TSFVLLAFFENHELIPKYQSAIKKAVRYVAEEADK-TDDQYSLAIAAVALQLAKHPQSEK 1082

Qy       1107 ALNMLTWRAEQEGGMQFWVSSESKLSDS---------WQPRSLDIEVAAYALLSHFLQFQ 1157
               :  |   | :|    :|     || ::|               |:||| |:|: :| ||:    :
Db       1083 VIAKLESVARKENDRMWW----SKATESTGEDGRVFHWKPRSNDVEITSYVLLALLEKDP 1138

Qy       1158 TSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFA--ALMNTERTNIQVTVTGPS-- 1213
                : :||::||  ||||  |||:|||||  : |:||::||          :    :|: :  |  |
Db       1139 AEKALPIIKWLISQRNSNGGFSSTQDTVIGLQALTKFAYKTGSGSGTMDIEFSSAGESKN 1198

Qy       1214 ----SPSPLAVVQ-------PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQE 1262
                    :|      |:|              |: :| | | |: ||:  ||:                :  :
Db       1199 TIKVNPENSLVLQTHDLPKSTRKVDFTAKGTGSAMVQLSYRYNLAEK---------EKKP 1249

Qy       1263 AFDLDVAVKENKDDLNHVDLNVCTSF------SGPGRSGMALMEVNLLSGFMVPSEA--- 1313
               :|     ||:   :    ||  || :              | ||:|:  | |||:  |   :
Db       1250 SFKVTPTVKDTPNQLLIVD--VCAEYVPLEDADKDKDSNMAVMEIALPSGFVGDSTSLGK 1307

Qy       1314 ISLSETVKKVEYDH--GKLNLYLDSVNETQF-CVNIPAVRNFKVSNTQDASVSIVDYYEP 1370
               |     : ||:||    :      : :| ||:          |: : |  :   |:  :  ||||: |||:
Db       1308 IQAVDRVKRVETKNSDSTVVVYFDSLTPGDVRCLPLEASKAHAVAKQKPASVSLYDYYDT 1367

Qy       1371 RRQAVRSYNSEVKLSSCDLCSDV---QGCR 1397
                  |:|    |   :|| |  ||:|          :||:
Db       1368 ERKATEYY--QVKSSLCDICEGADCGEGCK 1395


RESULT 9
Q8IPH4
ID   Q8IPH4      PRELIMINARY;      PRT;   1408 AA.
AC   Q8IPH4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG7052-PD.
GN   TEPII.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA   Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA   Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
```

```
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003618; AAN10640.1; -.
SQ   SEQUENCE   1408 AA;  156629 MW;  A60FC3A3943431F4 CRC64;

  Query Match             23.0%;  Score 1693;  DB 5;  Length 1408;
  Best Local Similarity   31.2%;  Pred. No. 2.9e-94;
  Matches  461;  Conservative  293;  Mismatches  541;  Indels  182;  Gaps   54;

Qy         19 ALAVAPGPRFLVTAPGIIRPGG--NVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEA 76
Db         14 ALLVNATGIYSVVGPGTLRSNSKYNVVVSVHKADG-PSQ--IKVSLNGPSYNET------ 64

Qy         77 EGVFEKGSFKTLTLPSLPLNSADEIYEL--------RVTGRTQDEILFSNSTRLSFETKR 128
Db         65 ---------KQIELP--PMSTQNVEFEVPKLATGNYNLSAEGVSGVVFKNSTKLNYADKK 113

Qy        129 ISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTS--LNILIKDPKSNLIQQWLSQQSDL 186
Db        114 PSVFVQTDKATYKPADLVQFRILFLDENTRPAKIEKPISVIIIDGAQNRIKQLSDVKLTK 173

Qy        187 GVISKTFQLSSHPILGDWSIQVQVN-DQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKH 245
Db        174 GVFSGELQLSEQPVLGTWKISVSVDGDNRETKSFEVDKYVLPKFEVIVDTPKAVVIADKV 233

Qy        246 LNGTITAKYTYGKPVKGDVTLT-------FLPLSFWGKKKNITKTFKINGSANFSFNDEE 298
Db        234 IKATIRAKYTYGKPVKGKATVSMERSYGYFGDLNANGNKQE--KTIDVDGKGHVEFD--- 288

Qy        299 MKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTT 358
Db        289 --IIHWAQRG--QYLP-----PIKLFAVVTEELTGNKQNATATVVLHQQRYSIEPYERPE 339

Qy        359 VLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKI 418
Db        340 HFEANKSFIYQVVVKNVDGSPVT--NSAKNVKIGF-DKSYSYFHEPSPK--------TRI 388

Qy        419 NYTVPQS----GTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK-SP-SKTYIQLKT 472
Db        389 NFEAPVNENGIATFNVRLPD-SDSRYYRIFASFDGSENTIGSISKFEPTPMSREPLKIQV 447

Qy        473 RDENIKVGSPFELVVSGNKRLKELSYMVVSRGQL-------VAVGKQNSTMFSLTPENSW 525
Db        448 NTKKPRLGEQVSFDVVSIEDLPYFVYTIVARGNVILSDYVDVPDGQKTYTV-KFTPTFSM 506

Qy        526 TPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDS 585
Db        507 VPKATIYVYYVVNN-DLQFEEKTIDFEKEFSNSIDVS-APTNAKPSEEVKLRIK-TDADS 563

Qy        586 IVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVF--QECGLWVLTD 643
Db        564 FVGLLGVDQSVLLLKSGNDLSQDDIFNSLNIYQTS---TPWMNGYGRYPGQTSGLVTLTN 620

Qy        644 AN--LTKDYIDGVYDNAE-YAERFMEENEGHIVDIHDFSLGSSP----------HVRKHF 690
Db        621 ANYPYNTEFPDYVEDDPEIYA---FENNLDALPPMP--AIANFPPDTGNTVQPVEIRKNF 675

Qy        691 PETWIWLDTNMGYRI-YQEFEVT--VPDSITSWVATGFVISEDLGLGLTTTPVELQAFQP 747
Db        676 ADVWIW--QSIGRSVGEEEFTLTKKIPDTITSWVVTGFSLNPTSGIALTKNPSKIRVFQP 733

Qy        748 FFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSD-KFDILMTSSE-----INA 801
Db        734 FFVSTNLPYSVKRGEVIAIPVVIFNYLDKTLDADVVMDNSDQEYEFTEATNEVLEKAIDE 793

Qy        802 TGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYS 861
Db        794 VRRVKRVTIPANSGKSVSFMIRPKNVGFTTLKITATSALAGDAIHQKLKVEPEGVTLFEN 853

Qy        862 QSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGC 921
Db        854 RAVFINLKDQPEMS--QSLDADIPNEVVPQSEFIEFSVVGDLLGPTLQNLDNLVRMPYGC 911

Qy        922 GEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYD 981
Db        912 GEQNMVNFVPNILVLKYLEVTGRKLPSVESKARKFLEIGYQRELTYKHDDGSYSAFGKSD 971

Qy        982 PSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG 1041
Db        972 ASGSTWLTAYVMRSFHQAGTYTDIDPKVITAGLDFLVSKQKESGEFPEVGKL----FDNA 1027

Qy       1042 NKSPVTLTAYIVTSLLGYRKYQPNID--VQESIHFLESEFSRGISDNYTLALITYALSSV 1099
Db       1028 NQNPLALTSFVLLAFFENHELIPKYQSAIKKAVRYVAEEADK-TDDQYSLAIAAVALQLA 1086

Qy       1100 GSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDS---------WQPRSLDIEVAAYALL 1150
Db       1087 KHPQSEKVIAKLESVARKENDRMWW----SKATESTGEDGRVFHWKPRSNDVEITSYVLL 1142

Qy       1151 SHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFA--ALMNTERTNIQVT 1208
Db       1143 ALLEKDPAEKALPIIKWLISQRNSNGGFSSTQDTVIGLQALTKFAYKTGSGSGTMDIEFS 1202

Qy       1209 VTGPS------SPSPLAVVQ-------PMAVNISANGFGFAICQLNVVYNVKASGSSRRR 1255
Db       1203 SAGESKNTIKVNPENSLVLQTHDLPKSTRKVDFTAKGTGSAMVQLSYRYNLAEK------ 1256

Qy       1256 RSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSF------SGPGRSGMALMEVNLLSGFMV 1309
Db       1257 ---EKKPSFKVTPTVKDTPNQLLIVD--VCAEYVPLEDADKDKDSNMAVMEIALPSGFVG 1311

Qy       1310 PSEA---ISLSETVKKVEYDH--GKLNLYLDSVNETQF-CVNIPAVRNFKVSNTQDASVS 1363
Db       1312 DSTSLGKIQAVDRVKRVETKNSDSTVVVYFDSLTPGDVRCLPLEASKAHAVAKQKPASVS 1371

Qy       1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDV---QGCR 1397
Db       1372 LYDYYDTERKATEYY--QVKSSLCDICEGADCGEGCK 1406


RESULT 10
Q9NFV7
ID   Q9NFV7      PRELIMINARY;      PRT;  1420 AA.
AC   Q9NFV7; Q9VLY8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   TEP2 protein precursor (CG7052 protein) (GH01829p).
GN   TEPII OR TEP2 OR CG7052 OR CG18589.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lagueux M., Perrodou E., Levashina E.A., Capovilla M., Hoffmann J.A.;
RT   "Constitutive expression of a novel complement-like protein in Toll
RT   and Jak gain-of-function mutants of Drosophila.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
```

```
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ269539; CAB87808.1; -.
DR   EMBL; AE003618; AAF52541.2; -.
DR   EMBL; AY122084; AAM52596.1; -.
DR   HSSP; P01024; 1C3D.
DR   FlyBase; FBgn0041182; TepII.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   InterPro; IPR000215; Serpin.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS00284; SERPIN; 1.
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
SQ   SEQUENCE   1420 AA;  158346 MW;  7713D5E1643F8A91 CRC64;

  Query Match             23.0%;  Score 1692;  DB 5;  Length 1420;
  Best Local Similarity   31.0%;  Pred. No. 3.4e-94;
  Matches  462;  Conservative  293;  Mismatches  539;  Indels  196;  Gaps   54;

Qy         19 ALAVAPGPRFLVTAPGIIRPGG--NVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEA 76
Db         14 ALLVNATGIYSVVGPGTLRSNSKYNVVVSVHKADG-PSQ--IKVSLNGPSYNET------ 64

Qy         77 EGVFEKGSFKTLTLPSLPLNSADEIYEL--------RVTGRTQDEILFSNSTRLSFETKR 128
Db         65 ---------KQIELP--PMSTQNVEFEVPKLATGNYNLSAEGVSGVVFKNSTKLNYADKK 113

Qy        129 ISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTS--LNILIKDPKSNLIQQWLSQQSDL 186
Db        114 PSVFVQTDKATYKPADLVQFRILFLDENTRPAKIEKPISVIIIDGAQNRIKQLSDVKLTK 173

Qy        187 GVISKTFQLSSHPILGDWSIQVQVN-DQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKH 245
Db        174 GVFSGELQLSEQPVLGTWKISVSVDGDNRETKSFEVDKYVLPKFEVIVDTPKAVVIADKV 233

Qy        246 LNGTITAKYTYGKPVKGDVTLT-------FLPLSFWGKKKNITKTFKINGSANFSFNDEE 298
Db        234 IKATIRAKYTYGKPVKGKATVSMERSYGYFGDLNANGNKQE--KTIDVDGKGHVEFD--- 288

Qy        299 MKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTT 358
Db        289 --IIHWAQRG--QYLP-----PIKLFAVVTEELTGNKQNATATVVLHQQRYSIEPYERPE 339

Qy        359 VLKPSLNFTATVKVTRADGNQLTLEERRNNVVIVTVTQRNYTEYWSGSNSGNQKMEAVQKI 418
Db        340 HFEANKSFIYQVVVKNVDGSPVT--NSAKNVKIGF-DKSYSYFHEPSPK--------TRI 388

Qy        419 NYTVPQS----GTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK-SP-SKTYIQLKT 472
Db        389 NFEAPVNENGIATFNVRLPD-SDSRYYRIFASFDGSENTIGSISKFEPTPMSREPLKIQV 447

Qy        473 RDENIKVGSPFELVVSGNKRLKELSYMVVSRGQL-------VAVGKQNSTMFSLTPENSW 525
Db        448 NTKKPRLGEQVSFDVVSIEDLPYFVYTIVARGNVILSDYVDVPDGQKTYTV-KFTPTFSM 506

Qy        526 TPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDS 585
Db        507 VPKATIYVYYVVNN-DLQFEEKTIDFEKEFSNSIDVS-APTNAKPSEEVKLRIK-TDADS 563

Qy        586 IVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVF--QECGLWVLTD 643
```

```
Db        564 FVGLLGVDQSVLLLKSGNDLSQDDIFNSLNIYQTS---TPWMNGYGRYPGQTSGLVTLTN 620

Qy        644 AN-------LTKDYI---------------DGVYDNAEYAERFMEE--NEGHIVDIHDFS 679
              ||       |  |:                |:  ::     |: :  |: :: :|
Db        621 ANYPYNTGPLVMSYVFEGSRHPWITRPRYRVGIRGDSGDRISFLSQSLNDRNLKEI---L 677

Qy        680 LGSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVT--VPDSITSWVATGFVISEDLGLG 734
              |  :|     :|| |||||  :    |:|     :|| :|  :||:||||| |||  ::    |:
Db        678 LKQTPQRTTIRKEFPETWFF--ENVG---EEEFTLTKKIPDTITSWVVTGFSLNPTSGIA 732

Qy        735 LTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSD-KFDIL 793
              ||  | :::  |||||:  |||||| |||   |:  : |||||      :    |::: || :::
Db        733 LTKNPSKIRVFQPFFVSTNLPYSVKRGEVIAIPVVIFNYLDKTLDADVVMDNSDQEYEFT 792

Qy        794 MTSSE-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQM 848
               ::|       |:        : : :|:   | :| | ||| ::|    : :|| |  | ||: |
Db        793 EATNEVLEKAIDEVRRVKRVTIPANSGKSVSFMIRPKNVGFTTLKITATSALAGDAIHQK 852

Qy        849 ILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSI 908
              : |: ||:     :::: ::| |      |  ::|      |     |  || :: : :||:|||::
Db        853 LKVEPEGVTLFENRAVFINLKDQPEMS--QSLDADIPNEVVPQSEFIEFSVVGDLLGPTL 910

Qy        909 NGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQ 968
                | :|:|||||||||||:|| ||| :| ||      :     ::: ||   |:  ||||| |:
Db        911 QNLDNLVRMPYGCGEQNMVNFVPNILVLKYLEVTGRKLPSVESKARKFLEIGYQRELTYK 970

Qy        969 REDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFW 1028
               :|||:||||   | ||||||:|:|:| | :|   | |||   |:       :|     || :|||
Db        971 HDDGSYSAFGKSDASGSTWLTAYVMRSFHQAGTYTDIDPKVITAGLDFLVSKQKESGEFP 1030

Qy       1029 DPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNID--VQESIHFLESEFSRGISDN 1086
              : |::        |::|: ||:::: :        |       :::: ::  |     :   |
Db       1031 EVGKL----FDNANQNPLALTSFVLLAFFENHELIPKYQSAIKKAVRYVAEEADK-TDDQ 1085

Qy       1087 YTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDS---------WQP 1137
              |:||:    ||      |:::: :   |    |  :|     :|      || ::|            |:|
Db       1086 YSLAIAAVALQLAKHPQSEKVIAKLESVARKENDRMWW----SKATESTGEDGRVFHWKP 1141

Qy       1138 RSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFA-- 1195
              || |:|: :| ||:    :      : :||::||   ||||  |||:||||| :  |:||::||
Db       1142 RSNDVEITSYVLLALLEKDPAEKALPIIKWLISQRNSNGGFSSTQDTVIGLQALTKFAYK 1201

Qy       1196 ALMNTERTNIQVTVTGPS------SPSPLAVVQ-------PMAVNISANGFGFAICQLNV 1242
                  :   :|: :  | |      |    |  |:|                   |: :| | | |: ||:
Db       1202 TGSGSGTMDIEFSSAGESKNTIKVNPENSLVLQTHDLPKSTRKVDFTAKGTGSAMVQLSY 1261

Qy       1243 VYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSF------SGPGRSGM 1296
               ||:                :  :  :|  :   ||:  :  |   ||  ||    :               | |
Db       1262 RYNLAEK---------EKKPSFKVTPTVKDTPNQLLIVD--VCAEYVPLEDADKDKDSNM 1310

Qy       1297 ALMEVNLLSGFMVPSEA---ISLSETVKKVEYDH--GKLNLYLDSVNETQF-CVNIPAVR 1350
              |:||:  | |||:  |  :    |     : ||:||   :     : :|  ||:          |: : |  :
Db       1311 AVMEIALPSGFVGDSTSLGKIQAVDRVKRVETKNSDSTVVVYFDSLTPGDVRCLPLEASK 1370

Qy       1351 NFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDV---QGCR 1397
                |:  : .||||: |||:  |:|     |  :||  | ||:|         :||:
Db       1371 AHAVAKQKPASVSLYDYYDTERKATEYY--QVKSSLCDICEGADCGEGCK 1418
```

RESULT 11
Q9VLY9

```
ID   Q9VLY9      PRELIMINARY;      PRT;  1399 AA.
AC   Q9VLY9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG7052 protein.
GN   TEPII OR CG7052 OR CG18589.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
```

```
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003618; AAF52539.2; -.
DR   HSSP; P01024; 1C3D.
DR   FlyBase; FBgn0041182; TepII.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; MacrogloblnA2.
DR   InterPro; IPR000215; Serpin.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS00284; SERPIN; 1.
SQ   SEQUENCE   1399 AA;  155783 MW;  3D8B7DF91EE39F59 CRC64;

  Query Match             23.0%;  Score 1688.5;  DB 5;  Length 1399;
  Best Local Similarity   31.3%;  Pred. No. 5.4e-94;
  Matches  461;  Conservative  285;  Mismatches  544;  Indels  183;  Gaps   52;

Qy         19 ALAVAPGPRFLVTAPGIIRPGG--NVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEA 76
Db         14 ALLVNATGIYSVVGPGTLRSNSKYNVVVSVHKADG-PSQ--IKVSLNGPSYNET------ 64

Qy         77 EGVFEKGSFKTLTLPSLPLNSADEIYEL--------RVTGRTQDEILFSNSTRLSFETKR 128
Db         65 ---------KQIELP--PMSTQNVEFEVPKLATGNYNLSAEGVSGVVFKNSTKLNYADKK 113

Qy        129 ISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTS--LNILIKDPKSNLIQQWLSQQSDL 186
Db        114 PSVFVQTDKATYKPADLVQFRILFLDENTRPAKIEKPISVIIIDGAQNRIKQLSDVKLTK 173

Qy        187 GVISKTFQLSSHPILGDWSIQVQVN-DQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKH 245
Db        174 GVFSGELQLSEQPVLGTWKISVSVDGDNRETKSFEVDKYVLPKFEVIVDTPKAVVIADKV 233

Qy        246 LNGTITAKYTYGKPVKGDVTLT-------FLPLSFWGKKKNITKTFKINGSANFSFNDEE 298
Db        234 IKATIRAKYTYGKPVKGKATVSMERSYGYFGDLNANGNKQE--KTIDVDGKGHVEFD--- 288

Qy        299 MKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTT 358
Db        289 --IIHWAQRG--QYLP-----PIKLFAVVTEELTGNKQNATATVVLHQQRYSIEPYERPE 339

Qy        359 VLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKI 418
Db        340 HFEANKSFIYQVVVKNVDGSPVT--NSAKNVKIGF-DKSYSYFHEPSPK--------TRI 388

Qy        419 NYTVPQS----GTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK-SP-SKTYIQLKT 472
Db        389 NFEAPVNENGIATFNVRLPD-SDSRYYRIFASFDGSENTIGSISKFEPTPMSREPLKIQV 447

Qy        473 RDENIKVGSPFELVVSGNKRLKELSYMVVSRGQL-------VAVGKQNSTMFSLTPENSW 525
Db        448 NTKKPRLGEQVSFDVVSIEDLPYFVYTIVARGNVILSDYVDVPDGQKTYTV-KFTPTFSM 506

Qy        526 TPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDS 585
Db        507 VPKATIYVYYVVNN-DLQFEEKTIDFEKEFSNSIDVS-APTNAKPSEEVKLRIK-TDADS 563

Qy        586 IVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVF--QECGLWVLTD 643
Db        564 FVGLLGVDQSVLLLKSGNDLSQDDIFNSLNIYQTS---TPWMNGYGRYPGQTSGLVTLTN 620

Qy        644 ANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHD--FSLGS-------SPHVRKHFPETW 694
Db        621 AN---------YPYNTERRIFIRPGIGFPRPLFNRVTVAGSLPPNVIPEPQVRKEFPENW 671

Qy        695 IW-LDTNMGYRIYQEFEVT--VPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIF 751
Db        672 IFNIFENVG---EEEFTLTKKIPDTITSWVVTGFSLNPTSGIALTKNPSKIRVFQPFFVS 728

Qy        752 LNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSD-KFDILMTSSE-----INATGHQ 805
Db        729 TNLPYSVKRGEVIAIPVVIFNYLDKTLDADVVMDNSDQEYEFTEATNEVLEKAIDEVRRV 788

Qy        806 QTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSIL 865
Db        789 KRVTIPANSGKSVSFMIRPKNVGFTTLKITATSALAGDAIHQKLKVEPEGVTLFENRAVF 848

Qy        866 LDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQN 925
Db        849 INLKDQPEMS--QSLDADIPNEVVPQSEFIEFSVVGDLLGPTLQNLDNLVRMPYGCGEQN 906

Qy        926 MINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGS 985
Db        907 MVNFVPNILVLKYLEVTGRKLPSVESKARKFLEIGYQRELTYKHDDGSYSAFGKSDASGS 966

Qy        986 TWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSP 1045
Db        967 TWLTAYVMRSFHQAGTYTDIDPKVITAGLDFLVSKQKESGEFPEVGKL----FDNANQNP 1022

Qy       1046 VTLTAYIVTSLLGYRKYQPNID--VQESIHFLESEFSRGISDNYTLALITYALSSVGSPK 1103
Db       1023 LALTSFVLLAFFENHELIPKYQSAIKKAVRYVAEEADK-TDDQYSLAIAAVALQLAKHPQ 1081

Qy       1104 AKEALNMLTWRAEQEGGMQFWVSSESKLSDS---------WQPRSLDIEVAAYALLSHFL 1154
Db       1082 SEKVIAKLESVARKENDRMWW----SKATESTGEDGRVFHWKPRSNDVEITSYVLLALLE 1137

Qy       1155 QFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFA--ALMNTERTNIQVTVTGP 1212
Db       1138 KDPAEKALPIIKWLISQRNSNGGFSSTQDTVIGLQALTKFAYKTGSGSGTMDIEFSSAGE 1197

Qy       1213 S------SPSPLAVVQ-------PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQ 1259
Db       1198 SKNTIKVNPENSLVLQTHDLPKSTRKVDFTAKGTGSAMVQLSYRYNLAEK---------E 1248

Qy       1260 NQEAFDLDVAVKENKDDLNHVDLNVCTSF------SGPGRSGMALMEVNLLSGFMVPSEA 1313
Db       1249 KKPSFKVTPTVKDTPNQLLIVD--VCAEYVPLEDADKDKDSNMAVMEIALPSGFVGDSTS 1306

Qy       1314 ---ISLSETVKKVEYDH--GKLNLYLDSVNETQF-CVNIPAVRNFKVSNTQDASVSIVDY 1367
Db       1307 LGKIQAVDRVKRVETKNSDSTVVVYFDSLTPGDVRCLPLEASKAHAVAKQKPASVSLYDY 1366

Qy       1368 YEPRRQAVRSYNSEVKLSSCDLCSDV---QGCR 1397
Db       1367 YDTERKATEYY--QVKSSLCDICEGADCGEGCK 1397

RESULT 12
O01717
ID   O01717      PRELIMINARY;      PRT;  1507 AA.
AC   O01717;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Alpha-2-macroglobulin.
OS   Limulus sp.
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC   Limulidae; Limulus.
OX   NCBI_TaxID=51645;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97175055; PubMed=9022715;
RA   Iwaki D., Kawabata S., Miura Y., Kato A., Armstrong P.B.,
RA   Quigley J.P., Nielsen K.L., Dolmer K., Sottrup-Jensen L., Iwanaga S.;
RT   "Molecular cloning of Limulus alpha-2-macroglobulin.";
RL   Eur. J. Biochem. 242:822-831(1996).
DR   EMBL; D83196; BAA19844.1; -.
DR   HSSP; P01024; 1C3D.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR002048; EF-hand.
DR   InterPro; IPR001599; MacrogloblnA2.
```

```
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS00018; EF_HAND; 1.
SQ   SEQUENCE   1507 AA;  168490 MW;  8EF4013495F63CED CRC64;

  Query Match             22.3%;  Score 1638;  DB 5;  Length 1507;
  Best Local Similarity   28.9%;  Pred. No. 7.4e-91;
  Matches  449;  Conservative  270;  Mismatches  559;  Indels  278;  Gaps   43;

Qy         28 FLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVS---------VLEAEG 78
              |::|||  :  ||  :   :  :  |  :              :||    |  :          |  |  |
Db         29 FILTAPKSLTPGKSNILNLHLFD------------IKTNGFLRIGVKDQDDGNVVAETEV 76

Qy         79 VFEK---GSFKTLTLPS-----LPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRIS 130
               |  |     |     ||:||         |     |:   |      :  :  ::  |        ::     ::
Db         77 SFNKDNPSSSIQLTIPSGVEVKRPKLYANGSY----SSPSSNDFFFEKD--INMHKDKLI 130

Qy        131 VFIQTDKALYKPKQEVKFRIVTLFSDFK--PYKTSLNILIKDPKSNLIQQWLSQQSDLGV 188
              ||:||||  ||||  |  ||  ||:      |   |    |  :|    :   |::|      ::   |        |:
Db        131 VFVQTDKPLYKPGQTVKVRILPTTPDLKLVPKETIGSFQIENPDGIVLGYWPMLSFAEGI 190

Qy        189 ISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLNG 248
                      |      |    |  |  |:   :  |      |::|:|  ||||||||  :    |    |     |:   :
Db        191 AQFELALPDEPTYGMWRIKGNIEDTEIYENFEVKEYVLPKFEVKITPPSYLLTNADSITW 250

Qy        249 TITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFK-----------INGSANFSFNDE 297
               |  |:||||:||:|                  :  :  :|:  :         |                :|  ||      ||::
Db        251 KICAQYTYGQPVEGTFVAETNVVKYNWEKEGVPVIHKEGLIDGCLDVTVNSSA-LGFNEQ 309

Qy        298 EMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYT 357
               :                          |     |  :    |||   |||   |  :  :::    :      |  :  :  |
Db        310 RL-----------------SYRAVNMFAEVTEKGTGIKMNATDSIYRTSNPLNIMYLEPT 352

Qy        358 T---VLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEA 414
              :       |||  |  |       :||  :  ||          |:            |  :  :         |  |:             :|:  |
Db        353 SGKGYLKPGLPFYGKLKVEKPDGTPAPGEQ------IELCRFADRERWNRKRWLEEKIRA 406

Qy        415 VQKINYTVPQSGTFKIEF-PILEDSSELQLKAYFL-------GSKSSMAVHSL----FKS 462
               ::    :|   ::|    |        |       |  :   :  ||     |                :|  :       ||          :  |
Db        407 CKE--FTSDEAGIIKFTVPPQTPDITSFRFKAKALQYGKKDGDNKLNQPQHSFTVSSWYS 464

Qy        463 PSKTYIQLKTRDENIKVGSPFEL---VVSGNKRLKELSYMVVSRGQLVAVGK-------- 511
              ||  :::||:     |  |:  |   |                :|  ::  ::     |  :::|   :|   |
Db        465 PSGSHLQLEPITEEIECGKPLTVKFKYTTGEEKKQKFYYQIMARNFIVDTGSFEHEFLLS 524

Qy        512 -------------QNSTMFSLTPEN--SW----------------------TPKACVIV 533
                                       :  |   ||  |  |          |                                             |   |   ::|
Db        525 EDKSGLTDETYLPIDVTALSLNPPNEPEWENNVIVPPHIGETSLTLIPSFEMNPSAKILV 584

Qy        534 YYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVD 593
              :|:  :|||   ::|     ||   |:        :||:   |   :   :   |         :|:::    :   |   ||   ||   |||
Db        585 FYVREDGETVADSTKITVKKCLRNKVGLKFGEEKVLPGASSTLQLTAS-PYSICGIGAVD 643

Qy        594 KSVNLMNASNDITMENVVHEL-------------------------------ELYNTGYY 622
              |||:::::   |   ||   |   |   ::|                                                        :::|:
Db        644 KSVHILSSDNRITEEEVFNKLGGHDYYWPKQATSDYKYCEDYKFKQTEGEHEGSFSSGFT 703

Qy        623 LGMFMNSFAVFQECGLWVLTDANL------TKDYIDGVYDNAEYAERFMEENEGHIV--- 673
                :::|       |   ||   ||  |::|    |                    :|         :|              |        :         :
Db        704 STNYLDSITAFDEAGLVVISDMELETRPCKPSGFEDGGRPCPQYDVAFAAPQAANRIGGG 763

Qy        674 -DIHDFSLG-----SSP--HVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGF 725
                :       |      |          :    |     :  |   :||||:|   :|    :    |:   :|   ||  |   :
Db        764 GEAGGFGGGIRKKTNKPVVEIRTYFPETWLWELQNIGATGELSLKRDIPHTITEWVGSAI 823

Qy        726 VISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIE 785
                |||:  |||::           ::  |||||:     ||||||||:   :  :|:||||   :       :|:   :   |
Db        824 CISEETGLGVSEA-ATVKGFQPFFVSFTLPYSVIRGEKVPIIVTVFNYLSECLPIKLSLE 882

Qy        786 KSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS------- 838
               :||||              |:       :     |     |                |     :   |:|     ||::  :||     |
Db        883 QSDKF-------EMQNDTNSYTSCVCGGKSDTTRWMIKPRSLGQVNLTVYGASLPNEAIC 935

Qy        839 -------PTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTG 891
                            ||  ||  |:  :||:   ||        |     :    |                  |         :     |        |        |  :   |
Db        936 GNQDYSTVTARDAATRQLLVEPEGFPKEDTWSTFACPKDQNGKFT-ATSDLLLPEDLVED 994

Qy        892 SERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKE 951
              |  |    ::     ||::||:|    |    |:|:|   ||||||:   |   |||::|||||          :||::||
Db        995 SARGYVSITGDLMGPAIKNLDHLVRLPTGCGEQNMVKFVPNIFVLDYLTATGSITDSIKE 1054

Qy        952 KALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLH 1011
              |||:   ||::||  |:     |:      |||:|||||    |    |:   :|:|||   |    |    :|:   :  :|    |::|    |:
Db       1055 KALNNMRKGYARQQNYRHPDGSYSAFGNRDKQGNLFLTAFVYRSFAQAERFILINKNKLN 1114

Qy       1012 RTYTWLKGHQKSNGEFWDPGRVIHSELQGG----NKSPVTLTAYIVTSLL--GYRKYQPN 1065
                   |    |:      |:|||     |     |::    :|     |:||               :::|      ||||::    |||    ||:
Db       1115 ETENWILNRQRSNGCFRKIGKLFNSALKGGISSNDETPAPLTAYVLISLLEAGYKN---- 1170

Qy       1066 IDVQESIHFLESEFSRGIS--------DNYTLALITYALSSVGSPKAKEALNMLTWRAEQ 1117
                                         |:      :|||                           |:|||       |||   |    |    |   |   ||:  |     |      ||
Db       1171 ---------ETVIDQGISCLEALSNPSTYSLALFAYATSLAGHPSAKDYLAKLEERAIT 1220

Qy       1118 EGGMQFWVSSESKLSDSWQPRSLDIEVAAYA---LLSHFLQFQTSEGIPIMRWLSRQRNS 1174
              |||     ||    |              :           |:   :|:|     ||          ||   |                ::           ||:|||::|:|
Db       1221 EGGKTFWKSPSS--GRYYWGNSIGVEIAGYAVLTLLQHGGASNLAKVTPIIRWLAKQQNY 1278

Qy       1175 LGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
              |||   |||||   :||:|:|:||    ::       :     :::|   |                              :   :   :
Db       1279 RGGFYSTQDTVIALQAMSKFATIIYKDELDLEVGVESSGFEKKIMLTKDNSILMQTFRLQ 1338

Qy       1223 --PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKEN--KDDLN 1278
                |   |:     |      |      |      |        :   |   ::   |||                            ::   |   |:|     ||     |   :|
Db       1339 TVPSPVDFEATGSGCGLVQTSLRYNVNTP---------PPRKGFHLEVTVKRGLYRDCIN 1389

Qy       1279 HVDLNVCTSFSGPGR-SGMALMEVNLLSGFMVPSEAIS-----LSETVKKVEYDHGKLNL 1332
                 :        |   :   |       |     |    ||:|:|:   ::||::       |:|                    :::   |   |        :|||
Db       1390 -AHIATCVKYDGKGGVSNMAVLEMKMVSGWIPDEESIKNIVDREELNLRRYEVDGNQLNL 1448

Qy       1333 YLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCD 1388
                |     :   :         |   |               ::  :|     |:   |::  :   ||||      ::   |    ||:   :               :|:
Db       1449 YFSELTDQNLCFNFWLEQDIEVQETKPATIRLYDYYELEQEVVTSYSID---ENCE 1501

RESULT 13
Q9ULD7
ID   Q9ULD7      PRELIMINARY;      PRT;  1884 AA.
AC   Q9ULD7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein KIAA1283 (Fragment).
GN   KIAA1283.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20039619; PubMed=10574462;
RA   Nagase T., Ishikawa K., Kikuno R., Hirosawa M., Nomura N., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes. XV.
RT   The complete sequences of 100 new cDNA clones from brain which code
RT   for large proteins in vitro.";
RL   DNA Res. 6:337-345(1999).
DR   EMBL; AB033109; BAA86597.1; -.
DR   HSSP; P01024; 1C3D.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR002350; kazal.
DR   InterPro; IPR001599; Macroglobln A2.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   Pfam; PF00050; kazal; 1.
DR   SMART; SM00280; KAZAL; 1.
```

```
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW   Hypothetical protein.
FT   NON_TER      1      1
SQ   SEQUENCE   1884 AA;  206524 MW;  D3C078F26B951D1A CRC64;

  Query Match             21.0%;  Score 1542.5;  DB 4;  Length 1884;
  Best Local Similarity   27.4%;  Pred. No. 7.1e-85;
  Matches  494;  Conservative  257;  Mismatches  555;  Indels  497;  Gaps   57;

Qy          5 PLLTAAHLLCVC----TAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              |||    ||        ||   |||   :|:  || :  | |      | |  :   :  | :|||:|
Db          7 PLLPLLLLLSARDGVRAAQPQAPG--YLIAAPSVFRAGVEEVISVTIF-NSPREVTVQA 63

Qy         61 ELLKTASNLTVSVLEAEG-VFEKGSFKTLTLPSLPLNSADEIYELRVTGR---TQDEILF 116
              :|:        |::::|  : :||: | | :|:        |       |:|  ||          ::  ||
Db         64 QLVAQGE----PVVQSQGAILDKGTIK-LKVPTGLRGQA----LLKVWGRGWQAEEGPLF 114

Qy        117 SNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLI 176
              | |  :: :  :    |||||||.:|:|:      |     ||:   :  :|      |     | ||: : :
Db        115 HNQTSVTVDGRGASVFIQTDKPVYRPQHRVLISIFTVSPNLRPVNEKLEAYILDPRGSRM 174

Qy        177 QQWLS-QQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQT 235
               :|    :       |: : :| ||   |:||:|  | |::      | :||:|  :||||||: :
Db        175 IEWRHLKPFCCGITNMSFPLSDQPVLGEWFIFVEMQGHAYNKSFEVQKYVLPKFELLIDP 234

Qy        236 PLYCSMNSKHLNGTITAKYTYGKPVKGDVT--LTFLPLSFWGKK--KNITKTFKINGSAN 291
              | |              ||:  |:||:||||  | :|   :|      : ::    :   : : :|  || || :
Db        235 PRYIQDLDACETGTVRARYTFGKPVAGALTINMTVNGVGYYSHEVGRPVLRTTKILGSQD 294

Qy        292 FSFNDEEMKNVMDSSNGLSEYLDLSSP----GPVEILTTVTESVTGISRNVSTNVFFKQH 347
              |     :|                 :      |      | | |  | || | :
Db        295 FDICVRDM------------IPADVPEHFRGRVSIWAMVT-SVDGSQQ---------- 329

Qy        348 DYIIEFFDYTTV---------------LKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
                :  | | |  |                        || |        |:::   ||:               |  :
Db        330 ---VAFDDSTPVQRQLVDIRYSKDTRKQFKPGLAYVGKVELSYPDGSP------AEGVTV 380

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQKIN-YT---VPQSGTFKIEFPILEDSSE---LQLK 444
               :                 | |    | | ||   | | |      | | :    |::     |: |
Db        381 QI----------------KAELTPKDNIYTSEVVSQRGLVGFEIPSIPTSAQHVWLETK 423

Qy        445 AYFLGSK--------SSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLK-E 495
                  | |          | ::: |  :  |||: |:||:      ::||        |
Db        424 VMALNGKPVGAQYLPSYLSLGSWY-SPSQCYLQLQPPSHPLQVGEEAYFSVKSTCPCNFT 482

Qy        496 LSYMVVSRGQLVAVGKQ----------------------------------------N 513
              | | |  :|| :|   |:|
Db        483 LYYEVAARGNIVLSGQQPAHTTQQRSKRAAPALEKPIRLTHLSETEPPPAPEAEVDVCVT 542

Qy        514 STMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEK 573
              |    ::|  |  |     ::|:|:  ::||  :  |:  :|   | |:|:: : : |    :  : |
Db        543 SLHLAVTP--SMVPLGRLLVFYVRENGEGVADSLQFAVETFFENQVSVTYSANETQPGEV 600

Qy        574 VSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTG------------ 620
               | |||      :   |  | :  |||||  |:  :      :|    |    |||  |:
Db        601 VDLRIRAAR-GSCVCVAAVDKSVYLLRSGFRLTPAQVFQELEDYDVSDSFGVSREDGPFW 659

Qy        621 ---------------YYLGMFMNSFAVFQECGLWVLTD-ANLTKDYIDGVYDNAEYAER 663
                            :   |:   :|   | | ||  |:||  :|        |:|   :  |
Db        660 WAGLTAQRRRRSSVFPWPWGITKDSGFAFTETGLVVMTDRVSLNHRQDGGLY--TDEAVP 717

Qy        664 FMEENEGHIVDIHDFSLGSSPH-------VRKHFPETWIWLDTNMGYRIYQ-EFEVTVPD 715
               : : | :|      ::    | |              |  ||||||    |:      :       | |||
Db        718 AFQPHTGSLV-----AVAPSRHPPRTEKRKRTFFPETWIWHCLNISDPSGEGTLSVKVPD 772

Qy        716 SITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLK 775
              ||||||       :|       |||:    |      |:  |:|||:    ||    :||||:    : ::::||:
Db        773 SITSWVGEAVALSTSQGLGI-AEPSLLKTFKPFFVDFMLPALIIRGEQVKIPLSVYNYMG 831

Qy        776 DATEV--KVIIEKSDKFDILMTSSEINATGH------QQTLLVPSEDGATVLFPIRPTHL 827
                 ||  |: : |  :|                    ||              : : |    :     :    :   |
Db        832 TCAEVYMKLSVPKGIQF-----------VGHPGKRHVTKKMCVAPGEAEPIWVVLSFSDL 880

Qy        828 GEIPITVTALS-------------------------PTASDAVTQMILVKAEGIEKSYSQ 862
              |   ||  ||:                             |    |  |  : ::|:|||: ::|:
Db        881 GLNNITAKALAYGDTNCCRDGRSSKHPEENHADRRVPIGVDHVRRSVMVEAEGVPRAYTY 940

Qy        863 SIL----------------------LDLT----------DNRL--------------- 873
              |                        | ||          | |:
Db        941 SAFFCPSERVHISTPNKYEFQYVQRPLRLTRFDVAVRAHNDARVALSSGPQDTAGMIEIV 1000

Qy        874 ----------------------QSTLKTLSF---------------------------- 882
                                    | | ||:
Db       1001 LGGHQNTRSWISTSKMGEPVASAHTAKILSWDEFRTFWISWRGGLIQVGHGPEPSNESVI 1060

Qy        883 ------------------------------------------SFPPNTVTGSERVQIT 898
                                                        |    : ||||     :
Db       1061 VAWTLPRPPEVQFIGFSTGWGSMGEFRIWRKMEVDESYSEAFTLGVPHGAIPGSERATAS 1120

Qy        899 AIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMR 958
               ||||:||::|  |  :|:|:|:|:||||||||:||||:::|  ||  |  :||:   :: :     ::
Db       1121 IIGDVMGPTLNHLNNLLRLPFGCGEQNMIHFAPNVFVLKYLQKTQQLSPEVERETTDYLV 1180

Qy        959 QGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLK 1018
              |||||:|  |:|:|||:||||   |||  |||:||||: | :|  :|  :|  :|     | |    :|:
Db       1181 QGYQRQLTYKRQDGSYSAFGERDASGSMWLTAFVLKSFAQARSFIFVDPRELAAAKSWII 1240

Qy       1019 GHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLL--GYRKYQPNIDVQESIHFLE 1076
                |:::|  |     |||::  ::|||     | ||||:|  :||  |              :  :: ||||
Db       1241 QQQQADGSFLAVGRVLNKDIQGGIHGIVPLTAYVVVALLETGTASEEERGSTDKARHFLE 1300

Qy       1077 SEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQ 1136
              |       :  |  |:  ||  ||||: :  ||  |  |||   |   |     |:    |       ||:|| 
Db       1301 SAAPLAM-DPYSCALTTYALTLLRSPAAPEALRKLRSLAIMRDGVTHW-----SLSNSWD 1354

Qy       1137 PR----------------SLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFAS 1180
                                | ::|:  |||||::  |       :    :|:::|||:|||:||||:|
Db       1355 VDKGTFLSFSDRVSQSVVSAEVEMTAYALLTYTLLGDVAAALPVVKWLSQQRNALGGFSS 1414

Qy       1181 TQDTTVALKALSEFAAL---------------------MNTERTNIQVTVTG--PSSPS 1216
              |||| |||:||:|:| |                        |||  :|   |    || |:
Db       1415 TQDTCVALQALAEYAILSYAGGINLTVSLASTNLDYQETFELHRTNQKVLQTAAIPSLPT 1474

Qy       1217 PLAVVQPMAVNISANGFGFAICQLNVVYNVK---ASGSSRRRRSIQNQE----------- 1262
              | |              ||| |   |   :  |::|  |||        |  :  :|  |
Db       1475 GLFV--------SAKGDGCCLMQIDVTYNVPDPVAKPAFQLLVSLQEPEAQGRPPPMPAS 1526

Qy       1263 -----------AFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPS 1311
                          |  |  |   ::  :   |  | |||  :   |  |  ||::||  |||||
Db       1527 AAEGSRGDWPPADDDDPAADQHHQEYK-VMLEVCTRWLHAGSSNMAVLEVPLLSGFRADI 1585

Qy       1312 EAISLSET-------VKKVEYDHGKLNLYLDSV-NETQFCVNIPAVRNFKVSNTQDASVS 1363
              |  || :         :|: |     ::  |  | : :      ||    |:|    |    |    ||
Db       1586 E--SLEQLLLDKHMGMKRYEVAGRRVLFYFDEIPSRCLTCVRFRALRECVVGRTSALPVS 1643

Qy       1364 IVDYYEPRRQAVRSYNSEV------KLSSCDLCSDVQ-------GCRPCEDGASGSHHS 1410
              : |||||  :|  | ||                     :|  :      |::|:              |  | | : :
Db       1644 VYDYYEPAFEATRFYNVSTHSPLARELCAGPACNEVERAPARGPGWFPGESGPAVAPEEG 1703

Qy       1411 SVI 1413
              : |
Db       1704 AAI 1706


RESULT 14
Q8IZJ3
ID   Q8IZJ3      PRELIMINARY;      PRT;  1885 AA.
AC   Q8IZJ3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Alpha-2 macroglobulin family protein VIP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Li Z., Wu X., Engvall E.;
RT   "Cloning and expression of VIP, a novel alpha 2 macroglobulin family
RT   member.";
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY101765; AAM50084.1; -.
SQ   SEQUENCE   1885 AA;  206655 MW;  3FB4A8551FBF326D CRC64;

  Query Match             21.0%;  Score 1542.5;  DB 4;  Length 1885;
  Best Local Similarity   27.4%;  Pred. No. 7.1e-85;
  Matches  494;  Conservative  257;  Mismatches  555;  Indels  497;  Gaps   57;

Qy          5 PLLTAAHLLCVC----TAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
Db          8 PLLPLLLLLLSARDGVRAAQPQAPG--YLIAAPSVFRAGVEEVISVTIF-NSPREVTVQA 64

Qy         61 ELLKTASNLTVSVLEAEG-VFEKGSFKTLTLPSLPLNSADEIYELRVTGR---TQDEILF 116
Db         65 QLVAQGE----PVVQSQGAILDKGTIK-LKVPTGLRGQA----LLKVWGRGWQAEEGPLF 115

Qy        117 SNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLI 176
Db        116 HNQTSVTVDGRGASVFIQTDKPVYRPQHRVLISIFTVSPNLRPVNEKLEAYILDPRGSRM 175

Qy        177 QQWLS-QQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQT 235
Db        176 IEWRHLKPFCCGITNMSFPLSDQPVLGEWFIFVEMQGHAYNKSFEVQKYVLPKFELLIDP 235

Qy        236 PLYCSMNSKHLNGTITAKYTYGKPVKGDVT--LTFLPLSFWGKK--KNITKTFKINGSAN 291
Db        236 PRYIQDLDACETGTVRARYTFGKPVAGALTINMTVNGVGYYSHEVGRPVLRTTKILGSQD 295

Qy        292 FSFNDEEMKNVMDSSNGLSEYLDLSSP----GPVEILTTVTESVTGISRNVSTNVFFKQH 347
Db        296 FDICVRDM-------------IPADVPEHFRGRVSIWAMVT-SVDGSQQ----------- 330

Qy        348 DYIIEFFDYTTV----------------LKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
Db        331 ---VAFDDSTPVQRQLVDIRYSKDTRKQFKPGLAYVGKVELSYPDGSP------AEGVTV 381

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQKIN-YT---VPQSGTFKIEFPILEDSSE---LQLK 444
Db        382 QI-----------------KAELTPKDNIYTSEVVSQRGLVGFEIPSIPTSAQHVWLETK 424

Qy        445 AYFLGSK--------SSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLK-E 495
Db        425 VMALNGKPVGAQYLPSYLSLGSWY-SPSQCYLQLQPPSHPLQVGEEAYFSVKSTCPCNFT 483

Qy        496 LSYMVVSRGQLVAVGKQ----------------------------------------N 513
Db        484 LYYEVAARGNIVLSGQQPAHTTQQRSKRAAPALEKPIRLTHLSETEPPPAPEAEVDVCVT 543

Qy        514 STMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEK 573
Db        544 SLHLAVTP--SMVPLGRLLVFYVRENGEGVADSLQFAVETFFENQVSVTYSANETQPGEV 601

Qy        574 VSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTG------------- 620
Db        602 VDLRIRAAR-GSCVCVAAVDKSVYLLRSGFRLTPAQVFQELEDYDVSDSFGVSREDGPFW 660

Qy        621 ----------------YYLGMFMNSFAVFQECGLWVLTD-ANLTKDYIDGVYDNAEYAER 663
Db        661 WAGLTAQRRRRSSVFPWPWGITKDSGFAFTETGLVVMTDRVSLNHRQDGGLY--TDEAVP 718

Qy        664 FMEENEGHIVDIHDFSLGSSPH-------VRKHFPETWIWLDTNMGYRIYQ-EFEVTVPD 715
Db        719 AFQPHTGSLV-----AVAPSRHPPRTEKRKRTFPETWIWHCLNISDPSGEGTLSVKVPD 773

Qy        716 SITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLK 775
Db        774 SITSWVGEAVALSTSQGLGI-AEPSLLKTFKPFFVDFMLPALIIRGEQVKIPLSVYNYMG 832

Qy        776 DATEV--KVIIEKSDKFDILMTSSEINATGH------QQTLLVPSEDGATVLFPIRPTHL 827
Db        833 TCAEVYMKLSVPKGIQF----------VGHPGKRHVTKKMCVAPGEAEPIWVVLSFSDL 881

Qy        828 GEIPITVTALS------------------------PTASDAVTQMILVKAEGIEKSYSQ 862
Db        882 GLNNITAKALAYGDTNCCRDGRSSKHPEENHADRRVPIGVDHVRRSVMVEAEGVPRAYTY 941

Qy        863 SIL----------------------LDLT----------DNRL--------------- 873
Db        942 SAFFCPSERVHISTPNKYEFQYVQRPLRLTRFDVAVRAHNDARVALSSGPQDTAGMIEIV 1001

Qy        874 ---------------------QSTLKTLSF---------------------------- 882
Db       1002 LGGHQNTRSWISTSKMGEPVASAHTAKILSWDEFRTFWISWRGGLIQVGHGPEPSNESVI 1061

Qy        883 --------------------------------------------SFPPNTVTGSERVQIT 898
Db       1062 VAWTLPRPPEVQFIGFSTGWGSMGEFRIWRKMEVDESYSEAFTLGVPHGAIPGSERATAS 1121

Qy        899 AIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMR 958
Db       1122 IIGDVMGPTLNHLNNLLRLPFGCGEQNMIHFAPNVFVLKYLQKTQQLSPEVERETTDYLV 1181

Qy        959 QGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLK 1018
Db       1182 QGYQRQLTYKRQDGSYSAFGERDASGSMWLTAFVLKSFAQARSFIFVDPRELAAAKSWII 1241

Qy       1019 GHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLL--GYRKYQPNIDVQESIHFLE 1076
Db       1242 QQQQADGSFLAVGRVLNKDIQGGIHGIVPLTAYVVVALLETGTASEEERGSTDKARHFLE 1301

Qy       1077 SEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQ 1136
Db       1302 SAAPLAM-DPYSCALTTYALTLLRSPAAPEALRKLRSLAIMRDGVTHW-----SLSNSWD 1355

Qy       1137 PR----------------SLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFAS 1180
Db       1356 VDKGTFLSFSDRVSQSVVSAEVEMTAYALLTYTLLGDVAAALPVVKWLSQQRNALGGFSS 1415

Qy       1181 TQDTTVALKALSEFAAL---------------------MNTERTNIQVTVTG--PSSPS 1216
Db       1416 TQDTCVALQALAEYAILSYAGGINLTVSLASTNLDYQETFELHRTNQKVLQTAAIPSLPT 1475

Qy       1217 PLAVVQPMAVNISANGFGFAICQLNVVYNVK---ASGSSRRRRSIQNQE----------- 1262
Db       1476 GLFV--------SAKGDGCCLMQIDVTYNVPDPVAKPAFQLLVSLQEPEAQGRPPPMPAS 1527

Qy       1263 -----------AFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPS 1311
Db       1528 AAEGSRGDWPPADDDDPAADQHQEYK-VMLEVCTRWLHAGSSNMAVLEVPLLSGFRADI 1586

Qy       1312 EAISLSET-------VKKVEYDHGKLNLYLDSV-NETQFCVNIPAVRNFKVSNTQDASVS 1363
Db       1587 E--SLEQLLLDKHMGMKRYEVAGRRVLFYFDEIPSRCLTCVRFRALRECVVGRTSALPVS 1644

Qy       1364 IVDYYEPRRQAVRSYNSEV------KLSSCDLCSDVQ-------GCRPCEDGASGSHHHS 1410
Db       1645 VYDYYEPAFEATRFYNVSTHSPLARELCAGPACNEVERAPARGPGWFPGESGPAVAPEEG 1704

Qy       1411 SVI 1413
Db       1705 AAI 1707

RESULT 15
Q60486
ID   Q60486      PRELIMINARY;      PRT;  1476 AA.
AC   Q60486;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Alpha-macroglobulin precursor.
```

```
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97164019; PubMed=9010766;
RA   Iwasaki H., Suzuki Y., Sinohara H.;
RT   "Cloning and sequencing of cDNAs encoding plasma alpha-macroglobulin
RT   and murinoglobulin from guinea pig: Implications for molecular
RT   evolution of alpha-macroglobulin family.";
RL   J. Biochem. 120:1167-1175(1996).
DR   EMBL; D84338; BAA12316.1; -.
DR   HSSP; P01023; 1BV8.
DR   InterPro; IPR002890; A2M_N.
DR   InterPro; IPR001599; Macroglobln A2.
DR   InterPro; IPR000531; TonB_boxC.
DR   Pfam; PF00207; A2M; 1.
DR   Pfam; PF01835; A2M_N; 1.
DR   PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
DR   PROSITE; PS00430; TONB_DEPENDENT_REC_1; 1.
KW   Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   CHAIN        24   1476       POTENTIAL.
SQ   SEQUENCE   1476 AA;  163614 MW;  0550693A068DC5C5 CRC64;

  Query Match             20.3%;  Score 1491;  DB 11;  Length 1476;
  Best Local Similarity   28.8%;  Pred. No. 6.7e-82;
  Matches  451;  Conservative  292;  Mismatches  524;  Indels  300;  Gaps   60;

Qy          4 PPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAEL- 62
              | ||    || :  |  :::  |:::|  | ::  |     | : ||       |||||  |
Db         11 PVLL----LLILLAADASISGKPQYMVLVPSLLHSGTPEKICL-LLTQLNETVTVKASLD 65

Qy         63 -LKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTR 121
               ::   :| ::::  : :|:  :|     |:|  |   |     : | |  |     | :
Db         66 TIRENGSLFMNMVAEKDLFQCVAF---TVPQSPYPEAVMFLTVEVEGPTHG---FRSRKT 119

Qy        122 LSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLN-----ILIKDPKSNLI 176
               :  ::|   ||:|||| :|||  | || |:|:|   :|:|      ||       | :||| | :
Db        120 VLVKSKDSLVFVQTDKPIYKPGQTVKARVVSLDENFRP----LNELFPLIFIQDPKGNRV 175

Qy        177 QQWLSQQSDLGVISKTFQLSSHPILGDWSIQV-QVNDQTYYQSFQVSEYVLPKFEVTLQT 235
              || : : : |: : :| ||| |:|| :|: | : :       : || |  |:|||||| :
Db        176 MQWQNLKLERGLTQLSFPLSSEPLLGSYSVVVHKESGGRMHHSFTVEEFVLPKFEVQVSM 235

Qy        236 PLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFL-----PLSFWGKKKNITKTF--KING 288
              |   :: :   :: ::|||||| ||:|:      :   ::  :   :    | |
Db        236 PKKITILEQEFTVSVCGRYTYGKPVPGNITMSICRNYNNPSACLSEE---SRAFCKKYNQ 292

Qy        289 SAN--------FSFNDEEMK-----------NVMDSSNGLSEYLDLSSPGPVEILTTVT 328
                |        || :::                : :     : |:  |  ||  |:|
Db        293 QLNSQGCFIQQVKTNDFQLRRKEYEMRLRVEAKIREEGTG----VQLTGTGFSEITATIT 348

Qy        329 ESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNN 388
              :                    : |    : ::| : |   |::   ||  :  : :
Db        349 K--------------------LSFVKVDSYVRPGVPFFGQVRL--VDGKNVPMPHKM-- 383

Qy        389 VVITVTQRNYTEYWSGSNSGNQKMEAVQ-KINYTVPQSGTFKIEFPILEDSS----ELQL 443
              : || :: ||      ||:  :   ||  |||  |   |:    : ::   :  |
Db        384 ITITASEANY-----HSNATTDENGLVQFSINTTNMIGTSLNIQVKHKDSTNCYDYQWLL 438

Qy        444 KAYFLGSKSSMAVHSLFKSPSKTYIQLKT--------RDENIKVGSPFELVVSGNKRLKE 495
              :|   |  :: || ||      |:::: |:     :  |        ::  |  :  |||
Db        439 EANEGASHTANAVFSL----SRSFVHLEPQLGKLPCHQTQTFKA----HYILKG-QELKE 489

Qy        496 L--SYMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDGEII 543
              |   |:::::| :|  |      :| :|     |:  |:   | |  |:: |  |  ||||
Db        490 LVFYYVIMAKGGIVQSGTYVLSVEQGNTKGHFSVSVPVESDLAPVARVLIYAILPSGEII 549

Qy        544 SDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNL----- 598
              :|  |  |:     ||:  |  :|: :: |: |  ||:: : | |:   :||:||| |
Db        550 ADSAKYNVENCLDNKVNLSFSEGQSLPASKTHLRVTAS-PQSLCALRAVDQSVLLRKPEA 608

Qy        599 -MNASN--------DIT---------MENVVHELELYNTGYYLGMP---------MNSFA 631
               ::||:        |:|          ||    : |||  |  |:            : :
Db        609 VLSASSVYALLPVKDLTGFPGLLGQQEENDGECVSLYNT-YIDGILYSPEPNINEKDMYG 667

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGSSPH------ 685
               :: || | |:  : |  :        |        :|        | | |
Db        668 FLKDMGLKVFTNTKIQKPQLCAHVQKFEVPTMAYSYSESS-------SFRSGPRRVPAVG 720

Qy        686 ------------VRKHFPETWIW---LDTNMGYRIYQEFEVTVPDSITSWVATGFVISED 730
                          || : ||||||    :   :  |     | |||||:|| |  |  | :| |
Db        721 IAATYSEPPKETVRTYSPETWIWDLKVTDSSG---VAEVEVTVPDTITEWKAGAFCLSND 777

Qy        731 LGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKF 790
              ||||: |    |:||||||: | :||||||||  |  |: |: |||  |     :  |  :| | ||
Db        778 TGLGLSPT-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPDCIRISVHLEASPKF 836

Qy        791 DILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTA-------------- 836
                  : |          |:: |      :  || :  :  || :   ||:|
Db        837 LAEPKAKE------QESYCVCGNERQTVSWVVTPKSLGNVNFTVSAEALESSELCGNEKT 890

Qy        837 LSPT--ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSER 894
               : ||     |  : : :||: |||||   : : |: ::|    :    : |     | | :  | |
Db        891 VVPTYGKKDTIIKPLLVEPEGIEKEETWTSLIRVSDTTVS---EKLHLELPSNVIQDSAR 947

Qy        895 VQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKAL 954
              :: :||:|| ::   : :|::||||||||||: ||||||:||||| : :||| ::| ||:
Db        948 ATVSILGDILGSAMQNIQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPDIKSKAI 1007

Qy        955 SFMRQGYQRELLYQREDGSFSAFG-NY-DPSGSTWLSAFVLRCFLEADPYIDIDQNVLHR 1012
              |::  ||||:| |:    |||:| || ||    |:|||:||||: | :|  || ||:  : :
Db       1008 SYLSTGYQRQLNYKHRDGSYSTFGENYRGGQGNTWLTAFVLKTFSQARKYIFIDEAHITQ 1067

Qy       1013 TYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESI 1072
                :||  || || ||   :::: ::|| : ::|:||| :|| :|| :  : ||    |: ::
Db       1068 ALSWLSQKQKDNGCFWSSGSLLNNAIKGGVEDEISLSAYITIALLEMSLPDTHPVVRNAL 1127

Qy       1073 HFLES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFWV 1125
                |||       | : | |   || ||: || :  |: : ||| |  :|     :|    |
Db       1128 FCLESAWKSAKEGTHG-SHVYTKALLAYAFALAGNQERKKEILKSLEDEGVKEDNSLHWA 1186

Qy       1126 SSES-KLSDSW----QPRSLDIEVAAYALLSHFLQFQTSEGIP----------IMRWLSR 1170
                :  |:|: :     |  | ::|: :| ||::     |:    |            |: |:::
Db       1187 RPQKPKVSEGFLFKSQAPSAEVEMTSYVLLAYL----TARPAPTPEDLTSATDIVNWVTK 1242

Qy       1171 QRNSLGGFASTQDTTVALKALSEFAA--LMNTERTNIQVTVTGPSSPSP--------LAV 1220
              |:|| ||::||||||  |||  |||::||        ||:    |||:     :  |      :
Db       1243 QQNSHGGYSSTQDTVVALHALSKYAAATFTRTEKA-AQVTIKSSGTFSTNFEVNHNNRLL 1301

Qy       1221 VQPMAV-------NISANGFGFAICQLNVVYNV-----------------KASGSSRRRR 1256
              :| :::          |:  | |   |||            :|   :
Db       1302 LQQVSLPTVSDSYTITVTGEGNVYLQTSLKYNVPSEKGTFPFALEAETVPQACDGPKAHT 1361

Qy       1257 SIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGF--MVPS-EA 1313
              | |                    :|||   |   |  ||:::| ::|||   :  |: :
Db       1362 SFQ--------------------ISLNVSYIGSRP-VSNMAIVDVKMVSGFIPLKPTVKN 1400

Qy       1314 ISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQ 1373
               :  || :  :  |   :  :|||  |:      ::    |:: :|  : :  |  : : ||||
Db       1401 LEKSEHISRTEVSNNHVLIYLDKVSNQTLSLSFFVVQDIEVRDLKPAIIKVYDYYETNEF 1460

Qy       1374 AVRSYNS 1380
              |:  |::
Db       1461 AIAEYHA 1467

Search completed: January 15, 2004, 18:14:03
Job time : 64 secs
```

```
                    GenCore version 5.1.6
            Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         January 17, 2004, 00:12:55 ; Search time 5629 Seconds
                                   (without alignments)
                                   6165.712 Million cell updates/sec

Title:          US-10-020-095-4
Perfect score:  7348
Sequence:       1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters:      45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10020095/runat_15012004_163822_18713/app_query.fasta_1.1607
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10020095_@CGN_1_1_4337_@runat_15012004_163822_18713 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      EST:*
                1:  em_estba:*
                2:  em_esthum:*
                3:  em_estin:*
                4:  em_estmu:*
                5:  em_estov:*
                6:  em_estpl:*
                7:  em_estro:*
                8:  em_htc:*
                9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  em_gss_hum:*
                18:  em_gss_inv:*
                19:  em_gss_pln:*
                20:  em_gss_vrt:*
                21:  em_gss_fun:*
                22:  em_gss_mam:*
                23:  em_gss_mus:*
                24:  em_gss_pro:*
                25:  em_gss_rod:*
                26:  em_gss_phg:*
                27:  em_gss_vrl:*
                28:  gb_gss1:*
                29:  gb_gss2:*

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	5343	72.7	4485	11	AK029247	AK029247 Mus muscu
	2	5304	72.2	4725	11	AK036799	AK036799 Mus muscu
	3	1480	20.1	4641	11	BC044072	BC044072 Xenopus l
c	4	1461	19.9	941	13	BX391819	BX391819 BX391819
	5	1406	19.1	1201	9	AL554092	AL554092 AL554092
	6	1405.5	19.1	4667	11	BC049868	BC049868 Mus muscu
	7	1280	17.4	879	13	BU517328	BU517328 AGENCOURT
c	8	1117	15.2	667	12	BQ028040	BQ028040 UI-H-CO0-
c	9	1022	13.9	1201	9	AL576908	AL576908 AL576908
	10	1007	13.7	1495	11	AK042169	AK042169 Mus muscu
	11	1006	13.7	978	13	BX386457	BX386457 BX386457
	12	979.5	13.3	836	10	BG173799	BG173799 602333853
c	13	972.5	13.2	638	13	BU624356	BU624356 UI-H-FG1-
	14	941	12.8	648	10	BB613975	BB613975 BB613975
	15	912	12.4	634	9	AA418644	AA418644 zv93f07.r
	16	902.5	12.3	613	14	CB158337	CB158337 K-EST0217
	17	895	12.2	919	13	BX370169	BX370169 BX370169
c	18	870	11.8	988	13	BX386456	BX386456 BX386456
	19	849.5	11.6	586	9	AA284110	AA284110 zc36a07.T
	20	833	11.3	473	13	BX279537	BX279537 BX279537
c	21	831	11.3	537	14	CA417651	CA417651 UI-H-FE0-
	22	830	11.3	515	13	BX472602	BX472602 DKFZp686C
c	23	829	11.3	544	10	BE143588	BE143588 MR0-HT016
	24	759	10.3	853	10	BE573834	BE573834 601331813
	25	750.5	10.2	835	14	CD242054	CD242054 AGENCOURT
	26	746	10.2	490	12	BI975931	BI975931 484508 MA
	27	738	10.0	437	9	AA132252	AA132252 zo06h07.r
c	28	732	10.0	456	9	AI278292	AI278292 ql78d07.x
	29	731	9.9	480	12	BM537249	BM537249 ha80a02.g
	30	692	9.4	684	9	AV245523	AV245523 AV245523
	31	687	9.3	806	12	BG919720	BG919720 602822407
c	32	671	9.1	722	14	CA420610	CA420610 UI-H-FG0-
	33	661	9.0	511	14	CA893496	CA893496 B0179D10-
	34	660	9.0	620	13	BQ266550	BQ266550 NISC_ff13
	35	649	8.8	512	14	CA566520	CA566520 K0402C12-
c	36	645	8.8	395	12	BG983552	BG983552 IL5-CN006
c	37	636	8.7	717	13	BQ775821	BQ775821 UI-H-FH0-
c	38	635	8.6	635	14	CA429478	CA429478 UI-H-FH1-
	39	633.5	8.6	673	10	BB639005	BB639005 BB639005
	40	626	8.5	476	13	BQ561040	BQ561040 H4067H10-
c	41	614	8.4	687	13	BQ773484	BQ773484 UI-H-FE0-
	42	612.5	8.3	432	14	CB158287	CB158287 K-EST0217
	43	610.5	8.3	819	13	BU434307	BU434307 603836006
	44	609	8.3	431	10	BE006130	BE006130 RC4-BN012
c	45	597	8.1	652	12	BQ020033	BQ020033 UI-H-ED0-

ALIGNMENTS

```
RESULT 1
AK029247
LOCUS       AK029247                4485 bp    mRNA    linear   HTC 05-DEC-2002
DEFINITION  Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
            library, clone:4831440K17 product:weakly similar to
            alpha-2-macroglobulin (fragments) [Limulus polyphemus], full insert
            sequence.
ACCESSION   AK029247
VERSION     AK029247.1  GI:26081244
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 4485)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..4485
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4831440K17"
                     /db_xref="taxon:10090"
                     /clone="4831440K17"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="0 day neonate"
     misc_feature    112..4439
                     /note="putative
                     weakly similar to alpha-2-macroglobulin (fragments)
                     [Limulus polyphemus] (PIR|A36260, evidence: FASTY,
                     51.8%ID, 76.7%length, match=168)"
BASE COUNT     1277 a    970 c   1010 g   1228 t
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:        4485
Score:                  5343.00        Matches:       1036
Percent Similarity:     83.70%         Conservative:  176
Best Local Similarity:  71.55%         Mismatches:    206
Query Match:            72.71%         Indels:        31
DB:                     11             Gaps:          7

US-10-020-095-4 (1-1428) x AK029247 (1-4485)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              |||:::         ||||||:::||||||||||||||||||:::|||      ||||||
Db        112 ATGCGCAGTCGGAGGCTCCTGAGCGCGGCCCACCTTCTCTGCTTGTGCGCGGTCGCGCTG 171

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              |||   |||||||||   |||||||||||||||||||||||||||||||||||||||   
Db        172 GCC---GCCCCAGGGTCCCGCTTTCTGGTGACAGCCCCAGGGATCATCAGACCCGGAGCA 228

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||:::|||||||||:::   |||   ||||||   |||||||||
Db        229 AATGTGACTATTGGGGTGGATCTCCTGGAAAATAGCCCCCCACAGGTCCTTGTAAAGGCT 288

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ::::::   |||   |||||||||   :::   |||:::|||||||||||||||||||||
Db        289 CAGGTGTTCAAGATAGCTTCCAACAAATCGAGATCCATCCTGGAAGCAGAAGGAGTCTTT 348

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
                 :::|||   ||||||||||||   ||||||:::|||||||||:::|||||||||:::
Db        349 CACAGAGGCCATTTCAAGACTCTCGTTCTTCCGGCACTACCTCTGAGCAGTGCAGATAAG 408

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||   :::   |||::::::::::::||||||:::|||||||||   |||
Db        409 ATTTATGAGCTACACATAAACGGACAATCAGAGAATGAGATCGTATTCTCCAACAGGACA 468

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||:::||||||:::|||   |||||||||   ||||||||||||||||||   |||
Db        469 CGTTTAACATTTGAGAGCAAGAGTATATCTGTCCTCATTCAGACAGACAAGGCATTCTAC 528

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              |||||||||||||||||||||||||||::::::||||||   ||||||   |||||||||
```

```
Db   529 AAGCCGAAGCAGGAAGTGAAGTTTCGGGTGCTCACACTCTGCTCAGATCTCAAGCCTTAT 588

Qy   161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
Db   589 AGGACATCCGTGGACATTTTCATTAAGGATCCGAAGTCTAATGTGATCCAACAGTGGTTT 648

Qy   181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
Db   649 TCTCAGAAAGGTGATCTTGGAGTCGTTTCCAAAACTTTTCAACTATCCTCTAATCCAATA 708

Qy   201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
Db   709 TTTGGTGACTGGTCCATTCAGGTTCAAGTGAATGATCAGCAATATTACCAGTCATTTCAG 768

Qy   221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
Db   769 GTCCTAGAATATGTGTTACCAAAGTTCGAAGTCACCGTGCAGACCCCGCTGTATTGTTCC 828

Qy   241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
Db   829 TTGAAGTCTAAACAGTTAAATGGCTCCGTCATAGCAAAGTATACATACGGGAAGCCAGTG 888

Qy   261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
Db   889 AAGGGAAGTCTGTCACTTACATTTTTACCTTTGTCTTTTGGGGAAAGAAGAAGAATATC 948

Qy   281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
Db   949 ACAAAAAGTTTTGAGATAAATGGATTTGCAAACTTCTCTTTTGATAATTATGAGATGAAA 1008

Qy   301 AsnValMet------------------AspSerSerAsnGlyLeuSerGluTyrLeuAsp 314
Db  1009 AAGGTAATGAATTTGAAACCGCTAACGGATGTCTCGGAAGGGAGTTATGAGAATGTGGAC 1068

Qy   315 LeuSerSerProGlyProValGluIleLeuThrThrValThrGluSerValThrGlyIle 334
Db  1069 CCGTCCTTTCCTGGGCCGGCAGAGATTATAGCTACTGTGACGGAATCACTCACAGGTATC 1128

Qy   335 SerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePhe 354
Db  1129 TCAAGAATGGCAAGCACCAATGTGTTTTTCAAGCAACATGATTACATCATTGAGATTTTT 1188

Qy   355 AspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLysValThrArg 374
Db  1189 GATTATACGACTGTCTTGAAACCATCTCTCAACTTCACAGCCACTGTGAAGGTCAGCCGC 1248

Qy   375 AlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIleThrValThr 394
Db  1249 TCTGATGGCAATCAACTAACTCCTGAAGAAATAGAAAATGATTTGGTCACAGTAGTGACC 1308

Qy   395 GlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAla 414
Db  1309 CAGGGAAAAATAACC---------ACCCT-GAGAGTCAGAGAGATCAGGAAATGGATTAT 1358

Qy   415 ValGlnLysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIle 434
Db  1359 ATCCAGACCGTAAATTATACTATCCCCCAAAATGGAATCATTAAGATTGAATTCCCAGTC 1418

Qy   435 LeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMet 454
Db  1419 ATGTCCATTTCCGGCGAGCTACAACTGAAGGCTTATTTTCTTGATGGTACAAGTTCTGTG 1478

Qy   455 AlaValHisSerLeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAsp 474
Db  1479 ACAGTCCACAGTATGTTCACCTCTCCTAGTAAGACGTACATCCAGCTTAAAACCAGAGAT 1538

Qy   475 GluAsnIleLysValGlySerProPheGluLeuValValSerGlyAsnLysArgLeuLys 494
Db  1539 GAATATATAAAGGTGGGGTCACCTTTTGATTTGATGGTTAGTGGCAACAGACAATTCAAG 1598

Qy   495 GluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSer 514
Db  1599 GACTTAAGCTATATGGTAATATCAAAGGGACAGTTGGTGGCTGCAGGCAAACAAAGTTCA 1658

Qy   515 ThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleValTyr 534
Db  1659 AGAACTTTCTCTTTAACACCAGAGGCTTCCTGGGCTCCAAAGGCCTGTATAATCGCATAT 1718

Qy   535 TyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGlnLeuVal 554
Db  1719 TATATTGCAGAGGATGGGGAAATTATAAATGATATTCTAAAAATCCCCGTTCAGCTTGTT 1778

Qy   555 PheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGluProSerGluLysVal 574
Db  1779 TTTGAAAATAAGGTAAAGCTGTTTTGGAGTAAACCTACAGTCAAGCCATCTGATAAGGTC 1838

Qy   575 SerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaValAspLys 594
Db  1839 TCTCTCAGGATCTCTGCGACACAGTCTGACTCCCTGGTGGGGATTGTAGCTGTTGACAAA 1898

Qy   595 SerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValValHisGluLeu 614
Db  1899 AGTGTAACATTGATGGAAAACTCCAACAGTATTACAATGGAAACCATGGTCCATGAACTG 1958

Qy   615 GluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGln 634
Db  1959 GAACTTTATAACACAGAATATTATCTAGGCATGTTCATGAACTCCTTCGCTGTATTTCAG 2018

Qy   635 GluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyVal 654
Db  2019 GAATGTGGTCTCTGGGTGTTGACAGATGCAACCCTTATAAGAGATAGCATTGATGAAGTC 2078

Qy   655 TyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluGlyHisIleValAsp 674
Db  2079 TACGATACTGAAGAGTATTCTGAACGCTTTGCAGAGGAAAATGAGGCAAACCTGGTTGAT 2138

Qy   675 IleHisAspPheSerLeuGlySerSerProHisValArgLysHisPheProGluThrTrp 694
Db  2139 TTTGAAGATGCTTCTTCAGTTAACAATGTACATGTCAGGAAGAATTTTCCAGAAACATGG 2198

Qy   695 IleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValPro 714
Db  2199 ATTTGGCTAGACGCCTACATGGGATCTAAGATTTATGAAGAATTTGAAGTTACTGTACCT 2258

Qy   715 AspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGly 734
Db  2259 GACTCCATCACTTCTTGGGTGGCTTCGGCTTTTGTCATCTCAGAAGACCTGGGTTTGGG 2318

Qy   735 LeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeu 754
Db  2319 TTAACAACCGTTCCAGCAGAGCTGCAAGCCTTCCAACCGTTTTTCCTTTTCCTGAATCTT 2378

Qy   755 ProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeu 774
Db  2379 CCGTACTCTGTTATCAGAGGTGAAGAGTTTGCTTTGGAAGTATCCATCGTCAATTATTTG 2438

Qy   775 LysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMet 794
Db  2439 AAAGATACAATTAAGGTTGTGATACTCATTGAGGAAAGTGACAGCTTTGATATTTTAATG 2498

Qy   795 ThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAsp 814
Db  2499 ACTTCAAATGACACCAATGGCACCATATACCGGAAGACTGTTCAAGTTCCCAGGGATAAT 2558

Qy   815 GlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIleThrVal 834
Db  2559 GGGGTCACTCTTGTTTTTCCCATCAAGCCCACGCACTTGGGAGAGATTCCCATCACAGTG 2618

Qy   835 ThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValLysAlaGlu 854
Db  2619 ACGGCTGCCTCACCCACTGCCTCTGATGCTGTCACCCAGACAATTGTAGTAAAGCCTGAA 2678

Qy   855 GlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGln 874
Db  2679 GGGATAGAAAAGTCGTATTCGAAATCTGTCTTATTGGATCTGACAGATAGCAACGTAGAA 2738
```

```
Qy        875 SerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArg 894
              |||    :::::::::  ||||||||||||||:::||||||   ||||||||||||
Db       2739 AGTAAGCAGCAATCTATGAGATTCTCCTTTCCTCCTGATACGGTCATTGGCAGTGAAAGA 2798

Qy        895 ValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeu 914
              ||||||||||||||||||||||||:::|||||  |||||||||||||:::||||||
Db       2799 GTTCAGATCACAGCAATTGGAGATATCCTTGGTTCCTCCATCAATGGCTTATCTTCACTG 2858

Qy        915 IleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyr 934
              |||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||
Db       2859 ATCCGGATGCCTTATGGATGTGGTGAACAGAACATGATATATTTGCTCCAAATATTTAC 2918

Qy        935 IleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeu 954
              |||||||||||||||||||||:::||||||||||||  ||||||||||||||||||
Db       2919 ATTTTGGATTATCTGACTAAACAGAAACAGCTGACAGTTAATTTAAAAGAAAAGGCCCTT 2978

Qy        955 SerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPhe 974
              |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2979 TCATATATGAGGCAAGGTTACCAAAGGGAGCTTCTCTATCAGAGGGAAGATGGCTCCTTC 3038

Qy        975 SerAlaPheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArg 994
              ||||||||||||:::    |||   ||||||||||||||||||||||||||||||||
Db       3039 AGTGCTTTTGGGGACATTGACTCTTCTGGGAGCACTTGGCTGTCAGCATTTGTTTTAAGA 3098

Qy        995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014
              ||||||||||||||||||   ||||||||||||||||:::||||||||||||||||||
Db       3099 TGCTTTCTGGAAGCTGATTACTATATAGATATTGATCAGGATGTGTTACACAGAACATAT 3158

Qy       1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034
              ||||||||     |||:::|||   |||||||||||||||:::|||||||||||||||
Db       3159 ACTTGGCTTAATGCACATAAGAAATTCAATGGTGAATTTTGGGAGCCAGGAAGAGTGATT 3218

Qy       1035 HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
              ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db       3219 CACAGTGAACTTCAAGGTGGCACCAAAAGCCCAGTAACACTTACGGCCTATATTGTGACT 3278

Qy       1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
              |||:::||||||||||:::||||||||||||||||||||||||:::||||||:::|||   |||
Db       3279 TCTGTCCTGGGATACAAAAAGTATCAGCCTAATATCGATGTACAAGACTCAATCAAGTTT 3338

Qy       1075 LeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyr 1094
              ||||||   ||||||||||||||||||||||||||||||||||||:::|||:::|||
Db       3339 TTGGAATTTGAATTCAGCAGAGGAATTTCGGACAATTATACCTTAGCAATTATATCCTAT 3398

Qy       1095 AlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArg 1114
              |||||||||:::||||||||||||||||||:::||||||||||||||:::|||       |||
Db       3399 GCCCTGTCCACAGTCGGGAGCCCTAAAGCAGAGGAGGCTTTGAACTTGCTGATGCAGCGA 3458

Qy       1115 AlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSerLysLeuSerAspSer 1134
              :::|||:::||||||       ||||||||:::||||||          ||||||   |||
Db       3459 TCAGAAAAGGAAGGAGACACACAGTTCTGGTTGTCATCAGGACCTGCGCTCTCTGGGTCC 3518

Qy       1135 TrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeu 1154
              ||||||||||||||||||:::||||||||||||||||||||||||||||||:::||||
Db       3519 TGGCAGCCGCGCTCCGTGGACATCGAAATCGCTGCTTATGCACTGCTGGCGCACACTCTG 3578

Qy       1155 GlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSer 1174
                          |||||||||||||||:::|||||||||||      :::||||||||||||
Db       3579 CAC---CATGTGTCTGAGGGAATCCCGGTTATGAGGTGGCTCATCCAGCAAAGAAACAGC 3635

Qy       1175 LeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPhe 1194
              ||||||||||||   |||||||||||||||   ||||||||||||||||||||||||||||
Db       3636 CTGGGAGGTTTTGTATCCACTCAGGATACTGTTGTGGCCCTGAAGGCATTATCTGAATTT 3695

Qy       1195 AlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSer 1214
              :::||||||::::::   |||   |||:::||||||:::||||||||||||||
Db       3696 TCAGCCCTTGTGCACAAGGAAAATACAGATATCCAACTGACCGTCACAGGGCCCGGCATC 3755

Qy       1215 Pro--------------------------------------------------SerPro 1217
              |||
Db       3756 CCCAGATCCATACACTTCCGAATTGACTCCCAGAACCTCTTTCTCCTTCACCAGGAAGAG 3815

Qy       1218 LeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIle 1237
              |||      :::   |||:::   ||||||:::||||||:::|||   |||||||||||
Db       3816 CTTCATGCGCTAGATCCCATCACCGTTAATGTTTCTGCGCATGGCTCGGGATTTGCTATT 3875

Qy       1238 CysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSer 1257
              ||||||||||||||   ||||||||||||   ||||||||||||:::|||||||||||||
Db       3876 TGCCAGCTTAATGTTGACTATAACGTGAAAGGTTCAGGTTCTTCTAAAAGACGGAGATCT 3935

Qy       1258 IleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLysAspAspLeu 1277
                 :::||||||||||   |||||||||||||||   |||      |||:::||||||:::
Db       3936 ACCGAAAATCAAGAAGTTTTTGATTTAGACGTCATTGTG---AATAATGAGGACGACATT 3992

Qy       1278 AsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAla 1297
              :::|||::::::||||||||||||||||||      |||      |||:::||||||
Db       3993 AGTCACCTGAATCTGAATGTGTGCACAAGTCACTTGGGTTCAGAGAGGACAGGCATGGTC 4052

Qy       1298 LeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSer 1317
              |||||||||||||||||||||||||||||||         |||::::::|||   ||||||
Db       4053 CTTATGGAAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATTCAATTCCTCTGAGT 4112

Qy       1318 GluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerVal 1337
              ||||||:::||||||||||||||||||:::||||||||||||||||||||||||||||||
Db       4113 GAGACCCTGAAGAAAGTGGAATATGATAACGGGAAACTCAACCTTTATTTAGATTCTGTA 4172

Qy       1338 AsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThr 1357
              ||||||:::|||||||||||||||||||||||||||||||:::|||||||||||
Db       4173 AATGAATCCCAGTTTTGTGTTAATATTCCCACTGTGAGAGACTACAAAGTTTCAAATATT 4232

Qy       1358 GlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaValArgSer 1377
              :::|||   |||||||||::::::||||||||||||||||||||||||||||||||||||
Db       4233 CGAGATGGTTCAGTATCTGTAATGGACTACTATGAACCAAGGAGACAGGCAGTGCGAAGC 4292

Qy       1378 TyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArg 1397
              ||||||::::::||||||||||||||||||   |||      |||         |||:::
Db       4293 TACAACACCCAGGTGAAGCTGTCCTCATGTTACCTCAGTCCAGACACC---AACTGCAAG 4349

Qy       1398 ProCysGluAspGlyAlaSerGlySerHisHisHisSerSerValIlePheIlePheCys 1417
                       ||||||||||:::   |||         ||||||   :::   :::||||||
Db       4350 TCACACACGGACGGAGCCACGGACTCCCTTCGACGTTCTTCGAGCCTTCTCGTCTTTTGC 4409

Qy       1418 PheLysLeuLeuTyrPheMetGlu 1425
                    ||||||||||||::::::
Db       4410 TCCGTCCTTCTGTACTTTGTGCAA 4433
```

RESULT 2
AK036799

```
LOCUS       AK036799                4725 bp    mRNA    linear   HTC 05-DEC-2002
DEFINITION  Mus musculus adult female vagina cDNA, RIKEN full-length enriched
            library, clone:9930012E15 product:weakly similar to
            alpha-2-macroglobulin (fragments) [Limulus polyphemus], full insert
            sequence.
ACCESSION   AK036799
VERSION     AK036799.1  GI:26085434
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
```

```
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 4725)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1. .4725
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9930012E15"
                     /db_xref="taxon:10090"
                     /clone="9930012E15"
                     /sex="female"
                     /tissue_type="vagina"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    97. .4429
                     /note="putative
                     weakly similar to alpha-2-macroglobulin (fragments)
                     [Limulus polyphemus] (PIR|A36260, evidence: FASTY,
                     51.8%ID, 76.7%length, match=168)"
     polyA_signal    4707. .4712
                     /note="putative"
     polyA_site      4725
                     /note="putative"
BASE COUNT     1334 a   1016 c   1066 g   1309 t
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:         4725
Score:                  5304.00        Matches:        1037
Percent Similarity:     83.49%         Conservative:   177
Best Local Similarity:  71.32%         Mismatches:     202
Query Match:            72.18%         Indels:         38
DB:                     11             Gaps:           8

US-10-020-095-4 (1-1428) x AK036799 (1-4725)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              |||:::            |||||:::||||||||||||||||:::|||         ||||||
Db         97 ATGCGCAGTCGGAGGCTCCTGAGCGCGGCCCACCTTCTCTGCTTGTGCGCGGTCGCGCTG 156

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              |||   ||||||||   ||||||||||||||||||||||||||||||||||||||
Db        157 GCC---GCCCCAGGGTCCCGCTTTCTGGTGACAGCCCCAGGGATCATCAGACCCGGAGCA 213

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              ||||||||||||||||||:::|||||||||:::    |||      ||||||  |||||||||
Db        214 AATGTGACTATTGGGGTGGATCTCCTGGAAAATAGCCCCCCACAGGTCCTTGTAAAGGCT 273

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              ::::::     |||  |||||||||     :::     |||:::||||||||||||||||||||||
Db        274 CAGGTGTTCAAGATAGCTTCCAACAAATCGAGATCCATCCTGGAAGCAGAAGGAGTCTTT 333

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
                 :::|||    ||||||||||||     ||||||:::|||||||||:::|||||||||:::
Db        334 CACAGAGGCCATTTCAAGACTCTCGTTCTTCCGGCACTACCTCTGAGCAGTGCAGATAAG 393

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||     :::     |||:::::::::::||||||:::|||||||||     |||
Db        394 ATTTATGAGCTACACATAAACGGACAATCAGAGAATGAGATCGTATTCTCCAACAGGACA 453

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||:::||||||:::|||     |||||||||     ||||||||||||||||||     |||
Db        454 CGTTTAACATTTGAGAGCAAGAGTATATCTGTCCTCATTCAGACAGACAAGGCATTCTAC 513

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              ||||||||||||||||||||||||||||:::::::||||||     ||||||     |||||||||
Db        514 AAGCCGAAGCAGGAAGTGAAGTTTCGGGTGCTCACACTCTGCTCAGATCTCAAGCCTTAT 573

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              :::||||||::::::|||     |||||||||||||||||||||||||:::||||||||||||
Db        574 AGGACATCCGTGGACATTTTCATTAAGGATCCGAAGTCTAATGTGATCCAACAGTGGTTT 633

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||:::     ||||||||||||:::||||||||||||||||||||||||:::||||||
Db        634 TCTCAGAAAGGTGATCTTGGAGTCGTTTCCAAAACTTTTCAACTATCCTCTAATCCAATA 693

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
                 ||||||||||||||||||||||||||||||||||||     ||||||||||||||||||
```

```
Db    694 TTTGGTGACTGGTCCATTCAGGTTCAAGTGAATGATCAGCAATATTACCAGTCATTTCAG 753

Qy    221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
Db    754 GTCCTAGAATATGTGTTACCAAAGTTCGAAGTCACCGTGCAGACCCCGCTGTATTGTTCC 813

Qy    241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLys-ProVa 260
Db    814 TTGAAGTCTAAACAGTTAAATGGCTCCGTCATAGCAAAGTATACATACGGGAAGGCCAGT 873

Qy    260 lLysGly-AspValThrLeuThrPheLeuProLeuSerPheTrp-GlyLysLysLysAsn 279
Db    874 GAAGGGAAGGTCTGTCACTTACATTTTACCTTGTCTTTTGGGGAAAGAAGAAGAAT 933

Qy    280 IleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMet 299
Db    934 ATCACAAAAAGTTTTGAGATAAATGGATTTGCAAACTTCTCTTTTGATAATTATGAGATG 993

Qy    300 LysAsnValMet-----------------AspSerSerAsnGlyLeuSerGluTyrLeu 313
Db    994 AAAAAGGTAATGAATTTGAAACCGCTAACGGATGTCTCGGAAGGGAGTTATGAGAATGTG 1053

Qy    314 AspLeuSerSerProGlyProValGluIleLeuThrThrValThrGluSerValThrGly 333
Db   1054 GACCCGTCCTTTCCTGGGCCGGCAGAGATTATAGCTACTGTGACGGAATCACTCACAGGT 1113

Qy    334 IleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhe 353
Db   1114 ATCTCAAGAATGGCAAGCACCAATGTGTTTTTCAAGCAACATGATTACATCATTGAGATT 1173

Qy    354 PheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLysValThr 373
Db   1174 TTTGATTATACGACTGTCTTGAAACCATCTCTCAACTTCACAGCCACTGTGAAGGTCAGC 1233

Qy    374 ArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIleThr-Va 393
Db   1234 CGCTCTGATGGCAATCAACTAACTCCTGAAGAAATAGAAAATGATTTGGTCACAGTAGCT 1293

Qy    393 lThrGlnArg------AsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLy 411
Db   1294 GACCCAGCGGAAAAATAACCACCCTGAG---------------AGTCAGAGAGATCAGGA 1338

Qy    411 sMetGluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGl 431
Db   1339 AATGGATTATATCCAGACCGTAAATTATACTATCCCCCAAAATGGAATCATTAAGATTGA 1398

Qy    431 uPheProIleLeuGluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLy 451
Db   1399 ATTCCCAGTCATGTCCATTTCCGGCGAGCTACAACTGAAGGCTTATTTTCTTGATGGTAC 1458

Qy    451 sSerSerMetAlaValHisSerLeuPheLysSerProSerLysThrTyrIleGlnLeuLy 471
Db   1459 AAGTTCTGTGACAGTCCACAGTATGTTCACCTCTCCTAGTAAGACGTACATCCAGCTTAA 1518

Qy    471 sThrArgAspGluAsnIleLysValGlySerProPheGluLeuValValSerGlyAsnLy 491
Db   1519 AACCAGAGATGAATATATAAAGGTGGGGTCACCTTTTGATTTGATGGTTAGTGGCAACAG 1578

Qy    491 sArgLeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLy 511
Db   1579 ACAATTCAAGGACTTAAGCTATATGGTAATATCAAAGGGACAGTTGGTGGCTGCAGGCAA 1638

Qy    511 sGlnAsnSerThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysVa 531
Db   1639 ACAAAGTTCAAGAACTTTCTCTTTAACACCAGAGGCTTCCTGGGCTCCAAAGGCCTGTAT 1698

Qy    531 lIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProVa 551
Db   1699 AATCGCATATTATATTGCAGAGGATGGGGAAATTATAAATGATATTCTAAAAATCCCCGT 1758

Qy    551 lGlnLeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGluProSe 571
Db   1759 TCAGCTTGTTTTTGAAAATAAGGTAAAGCTGTTTTGGAGTAAACCTACAGTCAAGCCATC 1818

Qy    571 rGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAl 591
Db   1819 TGATAAGGTCTCTCTCAGGATCTCTGCGACACAGTCTGACTCCCTGGTGGGGATTGTAGC 1878

Qy    591 aValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValVa 611
Db   1879 TGTTGACAAAAGTGTAACATTGATGGAAAACTCCAACAGTATTACAATGGAAACCATGGT 1938

Qy    611 lHisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAl 631
Db   1939 CCATGAACTGGAACTTTATAACACAGAATATTATCTAGGCATGTTCATGAACTCCTTCGC 1998

Qy    631 aValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIl 651
Db   1999 TGTATTTCAGGAATGTGGTCTCTGGGTGTTGACAGATGCAACCCTTATAAGAGATAGCAT 2058

Qy    651 eAspGlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluGlyHi 671
Db   2059 TGATGAAGTCTACGATACTGAAGAGTATTCTGAACGCTTTGCAGAGGAAAATGAGGCAAA 2118

Qy    671 sIleValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHisPhePr 691
Db   2119 CCTGGTTGATTTTGAAGATGCTTCTTCAGTTAACAATGTACATGTCAGGAAGAATTTTCC 2178

Qy    691 oGluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluVa 711
Db   2179 AGAAACATGGATTTGGCTAGACGCCTACATGGGATCTAAGATTTATGAAGAATTTGAAGT 2238

Qy    711 lThrValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLe 731
Db   2239 TACTGTACCTGACTCCATCACTTCTTGGGTGGCTTCGGCTTTTGTCATCTCAGAAGACCT 2298

Qy    731 uGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePh 751
Db   2299 GGGTTTTGGGTTAACAACCGTTCCAGCAGAGCTGCAAGCCTTCCAACCGTTTTTCCTTTT 2358

Qy    751 eLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePh 771
Db   2359 CCTGAATCTTCCGTACTCTGTTATCAGAGGTGAAGAGTTTGCTTTGGAAGTATCCATCGT 2418

Qy    771 eAsnTyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAs 791
Db   2419 CAATTATTTGAAAGATACAATTAAGGTTGTGATACTCATTGAGGAAAGTGACAGCTTTGA 2478

Qy    791 pIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValPr 811
Db   2479 TATTTTAATGACTTCAAATGACACCAATGGCACCATATACCGGAAGACTGTTCAAGTTCC 2538

Qy    811 oSerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIlePr 831
Db   2539 CAGGGATAATGGGGTCACTCTTGTTTTTCCCATCAAGCCCACGCACTTGGGAGAGATTCC 2598

Qy    831 oIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuVa 851
Db   2599 CATCACAGTGACGGCTGCCTCACCCACTGCCTCTGATGCTGTCACCCAGACAATTGTAGT 2658

Qy    851 lLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAs 871
Db   2659 AAAGCCTGAAGGGATAGAAAAGTCGTATTCGAAATCTGTCTTATTGGATCTGACAGATAG 2718

Qy    871 nArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGl 891
Db   2719 CAACGTAGAAAGTAAGCAGCAATCTATGAGATTCTCCTTTCCTCCTGATACGGTCATTGG 2778

Qy    891 ySerGluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLe 911
Db   2779 CAGTGAAAGAGTTCAGATCACAGCAATTGGAGATATCCTTGGTTCCTCCATCAATGGCTT 2838

Qy    911 uAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaPr 931
Db   2839 ATCTTCACTGATCCGGATGCCTTATGGATGTGGTGAACAGAACATGATATATTTGCTCC 2898
```

```
Qy        931 oAsnIleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGl 951
              ||||||||||||||||||||||||||||||||:::|||||||||||  ||||||||||
Db       2899 AAATATTTACATTTTGGATTATCTGACTAAACAGAAACAGCTGACAGTTAATTTAAAAGA 2958

Qy        951 uLysAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAs 971
              ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db       2959 AAAGGCCCTTTCATATATGAGGCAAGGTTACCAAAGGGAGCTTCTCTATCAGAGGGAAGA 3018

Qy        971 pGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPh 991
              ||||||||||||||||||||||:::  |||   ||||||||||||||||||||||||
Db       3019 TGGCTCCTTCAGTGCTTTTGGGGACATTGACTCTTCTGGGAGCACTTGGCTGTCAGCATT 3078

Qy        991 eValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHi 1011
              |||||||||||||||||||||||||||     ||||||||||||||||:::||||||||
Db       3079 TGTTTTAAGATGCTTTCTGGAAGCTGATTACTATATAGATATTGATCAGGATGTGTTACA 3138

Qy       1011 sArgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGl 1031
              ||||||||||||||||||        |||:::|||  |||||||||||||:::|||||
Db       3139 CAGAACATATACTTGGCTTAATGCACATAAGAAATTCAATGGTGAATTTTGGGAGCCAGG 3198

Qy       1031 yArgValIleHisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTy 1051
              |||||||||||||||||||||||||||||||      ||||||||||||||||||||||
Db       3199 AAGAGTGATTCACAGTGAACTTCAAGGTGGCACCAAAAGCCCAGTAACACTTACGGCCTA 3258

Qy       1051 rIleValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSe 1071
              ||||||||||||:::||||||||||:::|||||||||||||||||||||||||||:::||
Db       3259 TATTGTGACTTCTGTCCTGGGATACAAAAAGTATCAGCCTAATATCGATGTACAAGACTC 3318

Qy       1071 rIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLe 1091
              ||||   |||||||||   ||||||||||||||||||||||||||||||||||||||::
Db       3319 AATCAAGTTTTTGGAATTTGAATTCAGCAGAGGAATTTCGGACAATTATACCTTAGCAAT 3378

Qy       1091 uIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLe 1111
              :|||:::||||||||||||:::|||||||||||||||||:::||||||||||||:::||
Db       3379 TATATCCTATGCCCTGTCCACAGTCGGGAGCCCTAAAGCAGAGGAGGCTTTGAACTTGCT 3438

Qy       1111 uThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSerLysLe 1131
              |       |||:::|||:::||||||      |||||||||:::||||||      ||
Db       3439 GATGCAGCGATCAGAAAAGGAAGGAGACACACAGTTCTGGTTGTCATCAGGACCTGCGCT 3498

Qy       1131 uSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSe 1151
              ||||   ||||||||||||||||||:::||||||||||:::||||||||||||||||::
Db       3499 CTCTGGGTCCTGGCAGCCGCGCTCCGTGGACATCGAAATCGCTGCTTATGCACTGCTGGC 3558

Qy       1151 rHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGl 1171
              :|||   |||            ||||||||||||||:::||||||||||||    :::||
Db       3559 GCACACTCTGCAC---CATGTGTCTGAGGGAATCCCGGTTATGAGGTGGCTCATCCAGCA 3615

Qy       1171 nArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLe 1191
              ||||||||||||||||||||||    ||||||||||||||||   ||||||||||||||||
Db       3616 AAGAAACAGCCTGGGAGGTTTTGTATCCACTCAGGATACTGTTGTGGCCCTGAAGGCATT 3675

Qy       1191 uSerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGl 1211
              |||||||||:::||||||::::::   |||   |||:::|||||:::||||||||||||
Db       3676 ATCTGAATTTTCAGCCCTTGTGCACAAGGAAAATACAGATATCCAACTGACCGTCACAGG 3735

Qy       1211 yProSerSerPro--------------------------------------------- 1215
              ||||   |||
Db       3736 GCCCGGCATCCCCAGATCCATACACTTCCGAATTGACTCCCAGAACCTCTTTCTCCTTCA 3795

Qy       1216 ----SerProLeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGl 1234
                        |||      :::   |||:::  |||||:::||||||:::|||   ||
Db       3796 CCAGGAAGAGCTTCATGCGCTAGATCCCATCACCGTTAATGTTTCTGCGCATGGCTCGGG 3855

Qy       1234 yPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgAr 1254
              ||||||||||||||||||||||||    |||||||||||    |||||||||||:::||
Db       3856 ATTTGCTATTTGCCAGCTTAATGTTGACTATAACGTGAAAGGTTCAGGTTCTTCTAAAAG 3915

Qy       1254 gArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLy 1274
              ||||||||||    :::||||||||||    ||||||||||||||   |||    |||::
Db       3916 ACGGAGATCTACCGAAAATCAAGAAGTTTTTGATTTAGACGTCATTGTG---AATAATGA 3972

Qy       1274 sAspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSe 1294
              :||||||:::::|||:::::||||||||||||||||||      |||        |||::
Db       3973 GGACGACATTAGTCACCTGAATCTGAATGTGTGCACAAGTCACTTGGGTTCAGAGAGGAC 4032

Qy       1294 rGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIl 1314
              :|||||    ||||||||||||||||||||||||||||||            |||:::::||
Db       4033 AGGCATGGTCCTTATGGAAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATTCAAT 4092

Qy       1314 eSerLeuSerGluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLe 1334
              |    ||||||||||:::||||||||||||||:::||||||||||||||||||||||||
Db       4093 TCCTCTGAGTGAGACCCTGAAGAAAGTGGAATATGATAACGGGAAACTCAACCTTTATTT 4152

Qy       1334 uAspSerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysVa 1354
              ||||||||||||||:::||||||||||||||||||||||||     ||||||::::::||||
Db       4153 AGATTCTGTAAATGAATCCCAGTTTTGTGTTAATATTCCCACTGTGAGAGACTACAAAGT 4212

Qy       1354 lSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAl 1374
              ||||||    :::|||   ||||||||||::::::||||||||||||||||||||||||
Db       4213 TTCAAATATTCGAGATGGTTCAGTATCTGTAATGGACTACTATGAACCAAGGAGACAGGC 4272

Qy       1374 aValArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGl 1394
              ||||||||||||||||::::::|||||||||||||||||||   |||        |||
Db       4273 AGTGCGAAGCTACAACACCCAGGTGAAGCTGTCCTCGTGTTACCTCAGTCCAGACACC-- 4330

Qy       1394 nGlyCysArgProCysGluAspGlyAlaSerGlySerHisHisHisSerSerValIlePh 1414
                          |||:::      ||||||:::       |||         ||||||    :::
Db       4331 -AACTGCAAGTCACACACGGACGGAGCCACGGACTCCCTTCGACGTTCTTCGAGCCTTCT 4389

Qy       1414 eIlePheCysPheLysLeuLeuTyrPheMetGlu 1425
               :::||||||    ||||||||||||::::::
Db       4390 CGTCTTTTGCTCCGTCCTTCTGTACTTTGTGCAA 4423
```

```
RESULT 3
BC044072
LOCUS       BC044072                4641 bp    mRNA    linear   HTC 19-FEB-2003
DEFINITION  Xenopus laevis, clone IMAGE:4681494, mRNA.
ACCESSION   BC044072
VERSION     BC044072.1  GI:28422351
KEYWORDS    HTC.
SOURCE      Xenopus laevis (African clawed frog)
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 4641)
  AUTHORS   Klein,S. and Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-2003) National Institutes of Health, Xenopus Gene
            Collection (XGC), National Institute of Child Health and Human
            Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
            20892-7510, USA
  REMARK    NIH-MGC Project
COMMENT     Contact: XGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Igor Dawid
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Institute for Systems Biology
            http://www.systemsbiology.org
            contact: amadan@systemsbiology.org
            Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha
            Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 94 Row: b Column: 12
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein
            This clone has the following problem: frame shifted.
FEATURES             Location/Qualifiers
     source          1..4641
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
```

```
                   /db_xref="taxon:8355"
                   /clone="IMAGE:4681494"
                   /tissue_type="Embryo, stage 31/32, Xenopus"
                   /clone_lib="NICHD_XGC_Emb4"
                   /lab_host="DH10B"
                   /note="Vector: pCMV-SPORT6"
BASE COUNT     1322 a   1063 c   1147 g   1109 t
ORIGIN

Alignment Scores:
Pred. No.:              7e-147          Length:         4641
Score:                  1480.00         Matches:        433
Percent Similarity:     46.30%          Conservative:   287
Best Local Similarity:  27.85%          Mismatches:     564
Query Match:            20.14%          Indels:         273
DB:                     11              Gaps:           46

US-10-020-095-4 (1-1428) x BC044072 (1-4641)

Qy         49 LeuGluHisCysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsn 68
Db        226 CTGGAGGGGGCGAATGGAGAAAGCAAGATACAGCTCACCCTACAGATGGCCAAAAGTAAC 285

Qy         69 LeuThrVal------SerValLeuGluAlaGluGlyValPheGluLysGlySerPheLys 86
Db        286 ACAACTGTGATAGAGAAGACATCAGAGCAGAAATCCTACTTCAGATGTGTTTCATTTGAG 345

Qy         87 ThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyrGluLeuArgVal 106
Db        346 ---------GTGCCGCCT---CCCTCTGAAGAGAATGAAGAAGTGGCCACCATGCAGGTT 393

Qy        107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db        394 TCTATCCAGAGCGGAGGAGAGACCATA---AGCAACAGCAGTAAAGTTCTGGTGAAGAAA 450

Qy        127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGluVal 146
Db        451 GCAAGAACCAGTATCTTCGTTCAAACAGACAAAGCTATATACAAACCAGGACAGACAGTA 510

Qy        147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeu---Asn 165
Db        511 AGATTCCGAGTTGTTTTCCTTAAAGAAAACCTCCAACCAGAAGAAAGTCAGGTGCCAACC 570

Qy        166 IleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp 185
Db        571 ATCGAATTACAGGACCCAGGGAAGAACCGTATTGGCCAGTGGCTAAATGTGAGTCTCCAG 630

Qy        186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db        631 CAAGGAATCGCTGAGCTCTCACTTCCCCTCTCTTCCGAGCCTCCGTTGGGAGAATATTCC 690

Qy        206 IleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrVal 225
Db        691 ATTCATGTA------------AAGAACACTGTCCATGTCTTCAGTGTGGAAGAATATGTT 738

Qy        226 LeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSerLysHis 245
Db        739 CTGCCCAGATTTGAGGTCACTCTCCAACTCCCAAAAGTGGTTGTGGTTGTCACTGGCCAA 798

Qy        246 LeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAspValThr 265
Db        799 TTCCCTGTTAAAATCTGTGCCAGGTACACTTATGGAAAACCAGTGCAGGGA--------- 849

Qy        266 LeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIleThrLysThrPheLys 285
Db        850 ---------------------------------------------------ACCTATAAG 858

Qy        286 IleAsn---GlySerAlaAsnPheSerPhe---------------AsnAspGluGluMet 299
Db        859 GTCGATGTGTGCCGGAAGCATTACAGCTACAGAAGGGCATTTGGTAATGACAAAGGGAGA 918

Qy        300 LysAsnValMetAspSerSerAsnGlyLeu------------------------------ 309
Db        919 GATCTGTGTGCTGATTTCAGTGGGAAGCTGGATCAGTCTGGTTGCTCCACATTCGATGTG 978

Qy        310 ------SerGluTyrLeuAspLeuSerSerProGlyProValGluIle------LeuThr 325
Db        979 AAGTCTTCTAACTACAATCTTCTATCAAGCGAATGGCAGATGAACATTGAGGGGGCTGGC 1038

Qy        326 ThrValThrGluSerValThrGlyIleSerArgAsnValSerThrAsnValPhePheLys 345
Db       1039 TGGATTACAGAAGATGGGACAGGAATTCAGATTTCTGCAACTGATAAATCCCAAATTTCA 1098

Qy        346 GlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeuLysProSerLeuAsn 365
Db       1099 TATGACCTCACCAGTGTCACCTTCCTGGATGCTGATGATACTTATAAGCCTGGGATTCCA 1158

Qy        366 PheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeuThr------------ 381
Db       1159 TACATCGGGACAGTGAAGGTTGTAGATGCCAGTGACACCCCCATCCCCAAACATCCAGTG 1218

Qy        382 ---LeuGluGluArgArgAsnAsnValValIleThr---ValThrGlnArgAsnTyrThr 399
Db       1219 TATCTGATCGACACCGATGGAATAGTAAACATGACCTTGATGACCGATAAGAATGGACTT 1278

Qy        400 GluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsn 419
Db       1279 GCTCCCTTCAAGCTGGACAATACAGCAGATTGGAAGGCA--------------------- 1317

Qy        420 TyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGlu--------- 436
Db       1318 ---ACGGTGTCCCTGACGGCCAAGACACAGGTGAAGGAGCCTCTTATGCAGTATCAGTAC 1374

Qy        437 ---------------------AspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGly 449
Db       1375 GGCATGATGTTCCCTAGGCATGGGTTAGCGATGCTTAGACTCAACCGCTTTTATTCCCGC 1434

Qy        450 SerLysSerSerMetAlaValHisSerLeu---------PheLysSerProSerLysThr 466
Db       1435 AGCAAGAGTTTTTAAAGCTCCACTCCCAGAATGGGGAGTTTCTGTGTGAGGGGCAACAA 1494

Qy        467 TyrIleGlnLeuLysThrArgAspGluAsnIleLysValGlySerProPheGluLeuVal 486
Db       1495 GAAGTCCAGGTGGAATATCTTATCAAGCGCACTGAGTTGGGAGAG--------------- 1539

Qy        487 ValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetValValSerArgGlyGlnLeu 506
Db       1540 ---------GAGGCAAAGCACTTGGAGTTACACTACCTGGCAGTTTCTAAAGGTGCCATA 1590

Qy        507 ---------------------ValAlaValGlyLysGlnAsnSerThrMet--------- 516
Db       1591 AAAGAATCTGACGTTTTAGAAGTTCCAATAAATAATGATGGTGAAGTCCTGAGAGGCGAG 1650

Qy        517 ------PheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleValTyr 534
Db       1651 GTCTCCATCAAACTCAAGCTCAGTGTTGACGTCTCTCCGACTCTTCGAGTTCTGACCTAC 1710

Qy        535 TyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIleProValGlnLeuVal 554
Db       1711 GTTTACTCCCTGGTGGAGAGATGGTGGCAGATGCAGCAAAGTTCAAGCTACAAAAGTGC 1770

Qy        555 PheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGluProSerGluLysVal 574
Db       1771 TTTAGGAATAAGGTGTCAGTTGGCTTCTCCCCAGATGAAGTTCTACCAGGATCAGATGTT 1830

Qy        575 SerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaValAspLys 594
Db       1831 TCCCTACAAGTTCAGGCAGCT---GCCGGGTCTCTCTGTGGCCTGAGAGTGGTGGATAAG 1887

Qy        595 SerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnVal------------ 610
Db       1888 AGTGTGGTGTTAATGAAACCTGACCAAGAACTGACGGCCGACAAGATTCGGGGACTCTTC 1947

Qy        611 ---------------------------------ValHisGluLeuGlu------------ 615
Db       1948 CCAGTCAATGATTTTGGAAATTATGATTATCGGATTCAGGAAGCTGAGGATCACTGTCCA 2007
```

```
Qy    616 -----------------------------------------------LeuTyrAsnThr 619
Db   2008 TTCCGACCCTTTCTTCCTTGGCGACCATTCAATAGAATTTCTAGATCCATTATATATCCC 2067
Qy    620 GlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrp 639
Db   2068 CCATGGCATATGCCAGAGGAAGCTGATGTCTACGGCCTCTTCAAGAAAGTGCAACTAAAA 2127
Qy    640 ValLeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAsp--------- 656
Db   2128 ATCATCACCAGCGCTGACATTAAGAAACCAGTGGTGTGTCAACATTATGATTTTTTTAAG 2187
Qy    657 ------AsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluGlyHisIleValAsp 674
Db   2188 GCCAGGGCTACCAGGACTTCCAGGACCTACTGGAGACCGAATGATTG-CATTTGCTTCCC 2246
Qy    675 IleHisAspPheSerLeuGlySerSerProHis--------------------------- 685
Db   2247 CTGGGATA-TCTGCTGCCGGATCTATACCAAACCAGCTACCAGAAGCAGAGTCTGCAGCC 2305
Qy    686 ---------ValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMetGly 702
Db   2306 GTTGAAATTATCCGAACCTACTTTCCAGAGACGTGGATATGGGAGCTGGCATCTGTGGGC 2365
Qy    703 TyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrpValAla 722
Db   2366 GAGTCTGGTACAGCAGAGCTCCATCGCGCAGTCCCAGACACCATCACAGACTGGAATGCT 2425
Qy    723 ThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeu 742
Db   2426 GGAGCTTTCTGTATGGGACCCAGT---GGATTTGGCATC---TCCTCACCAACCTCCCTT 2479
Qy    743 GlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGlu 762
Db   2480 CGAGTCTTCCAGCCCTTCTTTGTGGAGCTCACTCTGCCGTACTCTGTGGTGCGAGGAGAG 2539
Qy    763 GluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysVal 782
Db   2540 TCATTCGCTCTCAAGGCCACTGTCTTTAACTACCTGAAGCAGTGTATAAAGGTTCAGACA 2599
Qy    783 IleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThr 802
Db   2600 GCCCTGGAGCCTTCACAAGAGTTGGAG------------------GAGGAACCATGTCCT 2641
Qy    803 GlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIle 822
Db   2642 GACTGTCAGTATAGCAGCTGTGTCTGTGCAGAGGAGAGCAAGACCTTCTACTGGAACGTT 2701
Qy    823 ArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSer------------ 838
Db   2702 AAAGCTTCTAACCTGGGAGAGGTGAATGTCACGGTAAAGACGGTGGCTCTGGATACTCAG 2761
Qy    839 ------------------------------------ProThrAlaSerAspAlaValThr 846
Db   2762 GACTTGTGCAATAATGAAATCCCTATAGTCCCTAAACAGGGGAGATCAGACACTGTGATT 2821
Qy    847 GlnMetIleLeuValLysAlaGluGlyIle------GluLysSerTyrSerGlnSerIle 864
Db   2822 AAGCCACTACTGGTGCAGCCAGGGGGAGTCCTGGAAGAAAAAGCTCACAGTTCTCTGCTC 2881
Qy    865 LeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPhe 884
Db   2882 TGCTGCCAAGCTGGTGATGATCATCCTAAAACA------GAGGAGATATCCCTTAAGGTC 2935
Qy    885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904
Db   2936 CCTGAAAATATTCTGAAGGATTCAGAAAGAGCATATGTGACGGTCCTTGGAGATCTGATG 2995
Qy    905 GlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGln 924
Db   2996 GGAACAGCCATGCAGAACCTGGACCGTCTCCTGGCCATGCCGTATGGATGTGGGGAGCAG 3055
```

```
Qy    925 AsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysLysGln 944
Db   3056 AACATGGTTCTCTTTGCCCCCAATATCTTCATTCTTCAGTATCTGGAAAAGACTCACCAA 3115
Qy    945 LeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTyrGlnArgGlu 964
Db   3116 CTGAGCCCTGAGATCCAGAGCAAAGCCAAAAGGTTCCTTGAAAGCGGGCTCCAGCGACAG 3175
Qy    965 LeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGly 984
Db   3176 CTGACATACAAACGGGATGACGGATCATACAGCGCATTTGGGCAAAGAGACAAAGAGGGG 3235
Qy    985 SerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAsp 1004
Db   3236 AACACATGGCTGACTGCATTTGTGGTGAAGTCCTTCAGCAAGGCTCGTCCCTACATCTTC 3295
Qy   1005 IleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsn 1024
Db   3296 ATAGATGAAAGTCACCTAAAACATTCTTTCACTTGGCTGAAGAACAATCAACATGAAAGT 3355
Qy   1025 GlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGlyAsnLysSer 1044
Db   3356 GGCTGGTTCCAGAGTGTGGGGAGACTCTTCAATAATGCTATGAAGGGAGGAATAGATGAT 3415
Qy   1045 ProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLysTyrGlnPro 1064
Db   3416 GAAACCTCTCTCTCTGCTTATGTAACAATGGCTTTGCTA---------------GAAGCT 3460
Qy   1065 AsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer------------ 1080
Db   3461 GGTCTATCTGTACAGGACCCAATG------GTTGAGGGTGCTTTATCCTGGCTGAGGAAG 3514
Qy   1081 ------ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSer 1098
Db   3515 GCTGCCAAAGATGTGAGCAGTATCTACACCCAGGCTCTGCTGGCCTACACTTTCACTCTG 3574
Qy   1099 ValGlySerProLysAlaLysGlu---AlaLeuAsnMetLeuThrTrpArgAlaGluGln 1117
Db   3575 AGTGGTGATACTGAACTCAGGGAGATGCTGCTGGCTAAACTGGATGAGAAGGCTGTAAAA 3634
Qy   1118 GluGlyGlyMetGlnPheTrp------------ValSerSerGluSerLysLeuSerAsp 1133
Db   3635 AGTGAAGGACAGTTACACTGGGAAAGGAAACCTGCCGCAGCTCCCTCTGATTTACCCTAC 3694
Qy   1134 SerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSerHisPhe 1153
Db   3695 TGGTACCAGGCTCCCTCAGCCGAGGTGGAGTTAACCTCCTATGTACTGTTGACCCTTCTC 3754
Qy   1154 ------LeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGln 1171
Db   3755 TCTGGCCTCAAAAAGGACCTGGGGAAAGCTTCTGAGATAGTCAACTGGCTGAGCAAGCAG 3814
Qy   1172 ArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeu 1191
Db   3815 CAAAATCCATATGGGGGATTCTCTTCTACTCAGGACACGGTTGTAGCTCTACAGGCTCTT 3874
Qy   1192 SerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGly 1211
Db   3875 GCTAAGTATGCCGAGGCCACATTCACTGACAAGGGAGATGTAACAGTCACAGTCAGTTCC 3934
Qy   1212 ProSer------------------------------------------------SerPro 1215
Db   3935 AAGACTGGATTCCATCAACAGTTCCATGTGGATCAAAATAACCGGCTCCTGCTACAGAAA 3994
Qy   1216 SerProLeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPhe 1235
Db   3995 TCCTCACTGTCAGATATC---CCTGGGGACTACTCTCTGTCAGCCAGTGGGAGCGGCTGT 4051
Qy   1236 AlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArg 1255
Db   4052 GTGTATGTACAGAGTGTCCTGAGATACAACATTCCCCCACCT------------------ 4093
Qy   1256 ArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLys--- 1274
```

```
                     :::::        |||  :::  |||          :::|||
Db       4094 ---------AGAAGTGATGCCACCTTCTCAGTACAAGTGAAGACACTACCAGACAAATGT 4144

Qy       1275 ------AspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyPro--- 1291
                    |||  |||    :::   |||   :::        ::::::|||
Db       4145 CCTCAAGATCTATTAAAATTCATGACGTTAGGCATCACTGTTAAATATACAGGAACTCGG 4204

Qy       1292 GlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
                 :::|||   |||||:::::|||      :::||||||||||||   :::||||||
Db       4205 GAGAAATCAAACATGGCTGTGATTGAGTGTAAGATGCTCTCTGGATTT---ATCCCATCA 4261

Qy       1312 GluAla------------IleSerLeuSerGluThrValLysLysValGluTyrAspHis 1327
                :::              :::      |||:::  :::|||:::  |||
Db       4262 AAGGGCTCAATCAGAGATCTGGAGAAGAGCAAAGCAATAAAGCGGAGTGAGATTCAGACA 4321

Qy       1328 GlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIlePro 1347
                    :::  ||||||||||||   :::                    :::
Db       4322 GACATGGTGACATTATATTTAGATGAGCTGGGACATGAGCCTCTCCATCTCGCCTTCACG 4381

Qy       1348 AlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1367
                    ::::::      |||   |||   :::   |||:::|||   :::   ||||||
Db       4382 GTGGAACAAGACATTGTAGTGAAGAACCTAAAGCCGGCCACGGTGAAAGTCTATGATTAT 4441

Qy       1368 TyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCys 1387
              ||||||          :::||||||   |||||||||
Db       4442 TATGAGACAGGAGAAGAGGCAGTTGCCGAGTACAACTCC--------------------- 4480

Qy       1388 AspLeuCysSerAspValGlnGlyCysArgProCys---GluAspGlyAlaSerGly--- 1405
                                                ||||||      ||||||      |||
Db       4481 ------------------------------CCCTGCAGCTCTGATGGAGAGAAAGGCAAC 4510

Qy       1406 ---------------------------------------------------SerHisHis 1408
                                                                 |||   :::
Db       4511 ACAAAATGAGATTGCTCCCGGCACTATTAACAGAAACATAGTTGAGCAAAGTCATCAAAC 4570

Qy       1409 HisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
                    |||:::   |||:::            ||||||      |||
Db       4571 AGATCTGCTGCAATATATTTTCACATAATGAAATTAAATATGTTT 4615
```

```
RESULT 4
BX391819/c
LOCUS       BX391819                 941 bp    mRNA    linear   EST 13-MAY-2003
DEFINITION  BX391819 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 3-PRIME, mRNA sequence.
ACCESSION   BX391819
VERSION     BX391819.1  GI:30627404
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 941)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAI029ZD07_CS02740_1&cluster=3578.f.
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0BAI029ZD07_CS02740_1.
FEATURES             Location/Qualifiers
     source          1..941
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      271 a    214 c    181 g    266 t      9 others
ORIGIN

Alignment Scores:
Pred. No.:              6.01e-146     Length:        941
Score:                  1461.00       Matches:       293
Percent Similarity:     93.95%        Conservative:  2
Best Local Similarity:  93.31%        Mismatches:    17
Query Match:            19.88%        Indels:        3
DB:                     13            Gaps:          0

US-10-020-095-4 (1-1428) x BX391819 (1-941)
```

```
Qy        941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
              |||      |||||||||      |||||||||       |||        ||||||
Db        938 AAAAGNAANCANCTGACAGATATTTGAAAAGAAAAGCTCTNTTCATTATGAGGGCAAGGG 879

Qy        961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
              |||||||||:::    |||||||||||||||||| ||||||||||||||||||||||||
Db        878 TACCAGAGAGACTTTCTCTATCAGAGGGAAGATGGCTC-TTCAGTGCTTTTGGGAATTAT 820

Qy        981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
              |||||||||||||||||||||||||||      |||   ||||||||||||||||||
Db        819 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTNTGTTNTAAGATGTTTCCTTGAAGCCGAT 760

Qy       1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 700

Qy       1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 640

Qy       1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        639 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 580

Qy       1061 LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        579 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 520

Qy       1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        519 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 460

Qy       1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGly 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        459 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 400

Qy       1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        399 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 340

Qy       1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        339 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 280

Qy       1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        279 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 220

Qy       1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        219 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 160

Qy       1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerProLeuAlaVal 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        159 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTG 100

Qy       1221 ValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeu 1240
              ||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         99 GTACAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTC 40

Qy       1241 Asn--ValValTyrAsnValLysAlaSerGlySerSer 1252
              |||    ||||||||||||:::   |||||||||||
Db         39 AATGNNTTGTATATAATGTAAGGCGTTCTGGGTCTTCT 2

RESULT 5
AL554092
LOCUS       AL554092                1201 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL554092 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 5-PRIME, mRNA sequence.
ACCESSION   AL554092
VERSION     AL554092.2  GI:31275905
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12894544.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI081DA09QP1&cluster=3578.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI081DA09QP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      359 a    238 c    236 g    293 t     75 others
ORIGIN

Alignment Scores:
Pred. No.:              7.4e-140      Length:         1201
Score:                  1406.00       Matches:        293
Percent Similarity:     90.74%        Conservative:   1
Best Local Similarity:  90.43%        Mismatches:     29
Query Match:            19.13%        Indels:         4
DB:                     9             Gaps:           1

US-10-020-095-4 (1-1428) x AL554092 (1-1201)

Qy          1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
              ||||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db         92 ATGCAGGGCCCACCGCTCCTGAMCGCMGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 151

Qy         21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
              |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db        152 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGAMAGCMCCAGGGATMATCAGGCCCGGAGGA 211

Qy         41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAla 60
              |||||||||||||||||||||||||||||||    ||||||        ||||||||||||
Db        212 AATGTGACTATTGGGGTGGAGCTTCTGGAAMACTGCCCTTMAMAGGTGACTGTGAAGGCG 271

Qy         61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80
              |||||||||||||||||||||        ||||||||||||||||||||||||||||||||
Db        272 GAGCTGCTCAAGACAGCATCAAAMMTCACTGTCTCTGTMCTGGAAGCAGAAGGAGTCTTT 331

Qy         81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
              ||||||  |||||||||  |||  ||||||        ||||||   ||||||||||
Db        332 GAAAAARGCTCTTTTAARAMACTTAMTCTTCCATMAMTACMTCTGAACARTGCAGATGAG 391

Qy        101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
              ||||||||||||   ||||||||||||||||||||||||||||||||||||    ||||||
Db        392 ATTTATGAGCTAMGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAAWAGTACC 451

Qy        121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140
              ||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db        452 CGCTTATCATTTGAGACCAAGAGAATATCTGTMTTCATTCAAACAGAMAAGGCCTTATAC 511

Qy        141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
              |||  |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 AAGCAAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 571

Qy        161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 631

Qy        181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        632 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 691

Qy        201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        692 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 751

Qy        221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        752 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 811

Qy        241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260
              |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        812 ATGAATTCTAAGCATTTAAATGGTACATC-ACGGCAAAGTATACATATGGGAAGCCAGTG 870

Qy        261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        871 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 930

Qy        281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              ||||||||||||||||||||||||||||||||||||||||     ||||||||||||
Db        931 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTWATGAWGAAGAGATGAAA 990

Qy        301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
                 ||||||||   ||||||||||||||||            :::        |||
Db        991 AW-GTWATGGATCT-TCAAATGGACTTTCTGAA---ACYKGGATYTATYTCCCCTGGACM 1045

Qy        321 ValGluIleLeu 324
              ||||||||||||
Db       1046 GTAGAAATTTTA 1057

RESULT 6
BC049868
LOCUS       BC049868                4667 bp    mRNA    linear   HTC 01-APR-2003
DEFINITION  Mus musculus, murinoglobulin 1, clone IMAGE:5097177, mRNA.
ACCESSION   BC049868
VERSION     BC049868.1  GI:29437204
KEYWORDS    HTC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4667)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
```

```
JOURNAL   Submitted (31-MAR-2003) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT   Contact: MGC help desk
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: Jeffrey E. Green, M.D.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC),
          Gaithersburg, Maryland;
          Web site: http://www.nisc.nih.gov/
          Contact: nisc_mgc@nhgri.nih.gov
          Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
          Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
          Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
          Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
          Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
          McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
          Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
          Young,A., Zhang,L.-H. and Green,E.D.

          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 99 Row: h Column: 5
          This clone was selected for full length sequencing because it
          passed the following selection criteria: matched mRNA gi: 6678965
          This clone has the following problem: frame shifted.
FEATURES           Location/Qualifiers
     source        1..4667
                   /organism="Mus musculus"
                   /mol_type="mRNA"
                   /strain="FVB/N"
                   /db_xref="taxon:10090"
                   /clone="IMAGE:5097177"
                   /tissue_type="Liver, normal. 5 month old male mouse."
                   /clone_lib="NCI_CGAP_Li9"
                   /lab_host="DH10B"
                   /note="Vector: pCMV-SPORT6"
BASE COUNT     1346 a   1126 c   1046 g   1149 t
ORIGIN

Alignment Scores:
Pred. No.:              7.39e-139     Length:          4667
Score:                  1405.50       Matches:         445
Percent Similarity:     45.10%        Conservative:    268
Best Local Similarity:  28.15%        Mismatches:      559
Query Match:            19.13%        Indels:          310
DB:                     11            Gaps:            53

US-10-020-095-4 (1-1428) x BC049868 (1-4667)

Qy          4 ProProLeuLeuThrAlaAlaHisLeu--------------------LeuCysValCys 16
              ||||||:::|||:::          :::                   ||||||:::
Db         73 CCTCCTGTCCTTTCTTACCACACCATGTGGAAGAGCAGACGGGCTCAGTTGTGCCTTTTT 132

Qy         17 ThrAlaAlaLeuAla--------------------ValAlaProGlyProArgPheLeu 29
              :::      ||||||                     :::          :::::::::
Db        133 TCGGTTCTCCTTGCCTTCCTGCATTCTGCTTCTTTACTCAACGGAGACTCGAAGTATATG 192

Qy         30 ValThrAlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeu 49
              |||      |||                                    :::|||
Db        193 GTGCTGGTCCCTTCT------------------------------------CAGCTCTAC 216

Qy         50 GluHisCysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeu 69
                       |||   ::::::   :::          |||                :::
Db        217 ACTGAGACCCCTGAGAAAATCTGCCTCCATCTATACCAGCTGAATGAAACAGTGACCGTC 276

Qy         70 ThrValSerValLeuGluAlaGluGly------------------------ValPheGlu 81
              |||   |||::::::         |||                        |||   :::
Db        277 ACAGCTTCCTTGGTATCTCAATCAGGAAGGAAAAACCTGTTCGATGAGCTGGTGCTTGAC 336

Qy         82 LysGlySerPheLysThrLeuThr------LeuProSerLeuProLeuAsnSerAlaAsp 99
              |||      |||:::   ::::::      :::|||   |||         :::   :::
Db        337 AAGGACTTGTTCCAATGTGTTTCCTTCATTATCCCTAGACTCTCCTCTTCTGATGAGGAG 396

Qy        100 GluIleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSer 119
              :::         :::   :::   |||   |||   :::         ||||||
Db        397 GATTTTCTCTACGTGGACATCAAAGGGCCAACCCACGAA---------TTCAGCAAAAGG 447

Qy        120 ThrArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeu 139
                 :::      :::   |||         ||||||:::||||||||||||   :::
Db        448 AAAGCAGTGCTTGTGAAGAACAAAGAAAGTGTTGTCTTTGTGCAGACAGATAAGCCCGTG 507

Qy        140 TyrLysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysPro 159
              |||||||||   |||   |||||||||||||:::|||::::::            :::|||
Db        508 TACAAGCCAGGACAATCAGTTAAATTTCGGGTTGTCTCTATGGATAAAATGCTACGTCCC 567

Qy        160 TyrLysThrSerLeuAsnIle---LeuIleLysAspProLysSerAsnLeuIleGlnGln 178
                       |||   :::      |||:::||||||||   |||   |||   |||
Db        568 CTGAATGAGTTGCTTCCTCTGGCTTACATTGAGGATCCGAAAAAGAACCGAATTATGCAG 627

Qy        179 TrpLeuSerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHis 198
              |||         :::::::::   |||:::         :::|||   |||::::::
Db        628 TGGAGGGATATTAAGACTGAGAATGGGCTTAAGCAAATGTCCTTCAGCCTGGCAGCAGAG 687

Qy        199 ProIleLeuGlyAspTrpSerIleGlnVal---GlnValAsnAspGlnThrTyrTyrGln 217
              ||||||   |||   :::   |||   |||   :::   :::   :::
Db        688 CCCATTCAGGGCCCCTACAAGATAGTGGTGCACAAAGAGTCAGGGGAGAAGGAAGAACAC 747

Qy        218 SerPheGlnValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeu 237
              ||||||   |||   |||:::|||||||||:::|||   |||   |||:::   |||
Db        748 TCCTTTACTGTGATGGAATTTGTGCTTCCCAGATTTAATGTCGACCTGAAGGTCCCAAAT 807

Qy        238 TyrCysSerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGly 257
                    |||:::|||   :::   |||:::   |||         |||||||||||||||
Db        808 GCCATGTCTGTGAATGATGAAGTGCTCAGTGTGACTGCATGTGGGAAATATACCTATGGG 867

Qy        258 LysProValLysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLys 277
              |||||||||   |||   |||
Db        868 AAGCCTGTTCCAGGACATGTG--------------------------------------- 888

Qy        278 LysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlu 297
                                   |||||||||
Db        889 ---------------------AAGATAAACGTCTGCCGTGAGACTGAGACTGGATGCAGA 927

Qy        298 GluMetLysAsnValMetAspSerSerAsnGlyLeuSer-----------GluTyrLeu 313
              |||:::   :::   :::|||   :::||||||   |||
Db        928 GAAGTCAATTCTCAGCTAGAC---AACAATGGCTGCAGTACACAAGAAGTGAACATCACT 984

Qy        314 AspLeuSerSerProGlyPro------------------ValGluIleLeuThrThrVal 327
              :::|||   |||                                  :::   ||||||
Db        985 GAGCTCCAATCAAAGAAAAGGAATTATGAGGTTCAGCTTTTCCATGTGAATGCAACTGTT 1044

Qy        328 ThrGluSerValThrGlyIle------SerArgAsnValSerThrAsnValPhePheLys 345
              ||||||      ||||||:::      ||||||:::   :::|||      :::
Db       1045 ACAGAAGAAGGGACAGGATTGGAGTTCAGTCGATCTGGAACAACTAAAATTGAAAGAATC 1104

Qy        346 GlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeuLysProSerLeuAsn 365
                 :::      |||      |||            :::      :::      :::
Db       1105 ACAAACAAGCTCATA------TTTCTGAAGGCAGATTCACACTTCAGACATGGAATTCCA 1158

Qy        366 PheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArg 385
              |||         |||::::::               |||:::   :::   |||:::
Db       1159 TTCTTTGTGAAAGTCCGCCTAGTGGATATCAAGGGAGACCCTATCCCAAATGAGAAA--- 1215

Qy        386 ArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySer 405
                       |||   |||         :::   :::||||||         :::
Db       1216 ---------GTCTTCATCAAAGCACAAGAACTTAGCTATACCAGTGCTACTACCACTGAT 1266

Qy        406 AsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyrThrValProGlnSer 425
                    |||
Db       1267 CAGCATGGC------------------------------------------------CTG 1278
```

```
Qy        426 GlyThrPheLysIleGluPheProIleLeuGluAspSerSerGluLeuGlnLeuLysAla 445
Db       1279 GCAGAGTTCTCCATAGATACCACATGCATCTCGGGCTCTTCC---CTCCATATCAAAGTC 1335
Qy        446 --------------------------------TyrPheLeuGlySerLysSerSerMet 454
Db       1336 AACCACAAAGAGGAAGATTCATGTTCCTATTTCTATTGCATGGAGGAAAGACATGCAAGT 1395
Qy        455 AlaValHisSerLeuPheLys------SerProSerLysThrTyrIleGlnLeuLysThr 472
Db       1396 GCAAAGCATGTGGCCTATGCTGTTTACTCCCTCAGCAAGAGCTACATCTACCTTGACACA 1455
Qy        473 ArgAspGluAsnIle---------------------LysValGlySerProPheGluLeu 485
Db       1456 GAGACCAGCAGCATCTTGCCCTGCAACCAGATTCACACAGTTCAGGCACATTTT------ 1509
Qy        486 ValValSerGlyAsnLysArg---LeuLysGluLeu------SerTyrMetValValSer 502
Db       1510 ATTCTGAAGGGGGACTTGGGAGTGCTGAAAGAGCTCATTTTCTACTACCTGGTCATGGCG 1569
Qy        503 ArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMet------------------ 516
Db       1570 CAGGGAAGCATCATCCAGACTGGAAACCATACTCACCAGGTGGAGCCAGGAGAAGCTCCA 1629
Qy        517 ------------------PheSerLeuThrProGluAsnSerTrpThrProLysAlaCys 530
Db       1630 GTAAAAGGAAAATTTGCCTTGGAGATCCCTGTGGAGTTTAGCATGGTCCCCATGGCTAAA 1689
Qy        531 ValIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIlePro 550
Db       1690 ATGCTCATCTACACGATCTTGCCTGATGGAGAAGTGATTGCAGATTCTGTAAACTTTGAA 1749
Qy        551 ValGlnLeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLysAlaGluPro 570
Db       1750 ATTGAAAAGTGTCTTCGCAACAAAGTGGACCTGCGCTTCAGCACTTCTCAAAGTCTTCCC 1809
Qy        571 SerGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleVal 590
Db       1810 GCCTCACAAACCCGTCTGCAGGTCACAGCTTCT---CCTCAGTCCCTCTGTGGCCTGAGA 1866
Qy        591 AlaValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnVal 610
Db       1867 GCTGTGGACCAGAGTGTGCTACTCCTGAAACCCGAGTCTGAGCTCTCCCCTTCC------ 1920
Qy        611 ValHisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSerPhe 630
Db       1921 ------------TGGATATACAATCTGCCA---------GGTATGCAGCAAAACAAATTC 1959
Qy        631 AlaValPhe---------------------GlnGluCysGlyLeu---------TrpVal 640
Db       1960 GTTCCAAGTTCCCGTCTGTCTGAAGACCAAGAAGACTGTATACTGTACAGTTCATGGTTG 2019
Qy        641 ---------------------------------------------------------Leu 641
Db       2020 GCTGAGAAGCACACAAACTTAGTACCACATGGAACTGAGAAGGATGTCTATAGATATGTG 2079
Qy        642 ThrAspAlaAsnLeuThrLys--------------------------------------- 648
Db       2080 GAGGACATGGGTTTAACAGCATTCACCAACTTGATGATCAAACTTCCTATAATTTGTTTT 2139
Qy        649 AspTyr------------------------------------------------------ 650
Db       2140 GACTATGGCATGGTTCCAATATCAGCACCTCGTGTGGAATTTGATCTTGCATTTACCCCT 2199
Qy        651 ---IleAsp-GlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGl 669
Db       2200 GAAATTTCTTGGAGTT---------------TACGCACAACATTGAGTAAAAGACCTGAA 2244
Qy        669 uGlyHisIleValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHi 689
Db       2245 GAGCCACCACGTAAAGATCCATCATCGAATGATCCA-CTTACAGAGACCATTCGAAAATA 2303
Qy        689 sPheProGluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPh 709
Db       2304 TTTTCCTGAAACCTGGGTCTGGGATATAGTCACAGTAAACTCCACAGGACTGGCTGAAGT 2363
Qy        709 eGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGl 729
Db       2364 GGAAATGACTGTCCCTGACACCATCACTGAATGGAAGGCAGGAGCTCTCTGCTTGTCCAA 2423
Qy        729 uAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePh 749
Db       2424 TGACACTGGCCTTGGCCTCTCTTCTGTG---GTACCTCTCCAAGCCTTCAAGCCCTTCTT 2480
Qy        749 eIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleTh 769
Db       2481 TGTGGAGGTCTCATTGCCCTATTCTGTGGTCCGTGGAGAAGCCTTCATGCTCAAGGCCAC 2540
Qy        769 rIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLy 789
Db       2541 TGTGATGAACTATCTCCCCACAAGCATGCAGATGAGTGTGCAGCTGGAAGCCTCTCCTGA 2600
Qy        789 sPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLe 809
Db       2601 TTTCACAGCTGTCCCAGTGGGAGACGAC------------------CAAGATTCTTACTG 2642
Qy        809 uValProSerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGl 829
Db       2643 CCTCAGTGCCAATGGGAGGCACACCTCATCCTGGCTGGTAACTCCCAAGTCTTTAGGGAA 2702
Qy        829 uIleProIleThrValThrAla---------------LeuSerProThrAlaSerAspAl 844
Db       2703 TGTGAATTTCTCTGTGTCTGCGGAAGCACAACAGTCCTCAGAGCCCTGTGGTTCTGAGGT 2762
Qy        844 aValThr-----------------------------Gln-MetIleLeuValLysAlaG 854
Db       2763 GGCCACAGTTCCTGAAACTGGGAGAAAGGACAGTAGTCAAAGTCCTGATAGTTGAGCCTG 2822
Qy        854 luGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuG 874
Db       2823 AAGGAATCAAGCAAGAGCATACCTTCAGCTCACTGTTCTGTGCATCAGATGCTGAGATAT 2882
Qy        874 lnSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluA 894
Db       2883 CT---------GAAAAAATGTCCCTGGTCCTCCCACCAACAGTAGTGAAAGATTCAGCAA 2933
Qy        894 rgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerL 914
Db       2934 GAGCCCATTTCTCTGTGATGGGTGATATCTTGAGTTCAGCCATAAGAAACACACAAAACC 2993
Qy        914 euIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleT 934
Db       2994 TTCTCCATATGCCCTATGGCTGTGGGGAGCAGAACATGGTCCTTTTGCTCCCAACATCT 3053
Qy        934 yrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaL 954
Db       3054 ACGTACTGAAATATCTGAATGAAACCCAACAGCTGACTCAGAAGATCAAGACTAAAGCCC 3113
Qy        954 euSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerP 974
Db       3114 TTGGGTTCCTCAGAGCTGGTTATCAGAGAGAGCTGAACTACAAACACAAGGATGGCTCCT 3173
Qy        974 heSerAlaPheGly------AsnTyrAspProSerGlySerThrTrpLeuSerAlaPheV 992
Db       3174 ACAGTGCCTTCGGGGATCAAAATGGCGAAAGAGAAGGAAACACTTGGCTCACAGCCTTTG 3233
Qy        992 alLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisA 1012
Db       3234 TGCTCAAGTCTTTTGCCCAAGCTCGAGCCTTCATCTTCATCGATGAATCACACATCACCC 3293
Qy       1012 rgThrTyrThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyA 1032
Db       3294 ATGCCTTCACCTGGCTCTCCCAAAAGCAGAAGGACAATGGCTGCTTCCGGAGCTCCGGAT 3353
Qy       1032 rgValIleHisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrI 1052
```

```
Db       3354 CATTGTTCAACAATGCCATGAAGGGGGGAGTAGATGATGAAATGACCCTCTCTGCCTACA 3413

Qy       1052 leValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerI 1072
              ||       :::||||||                        :::      |||   :::::::
Db       3414 TAACCATGGCCCTTCTGGAAAGTTCACTCCCAGCCACGCATCCTGTTGTCTCCAAAGCCC 3473

Qy       1072 leHisPheLeuGluSer---------------GluPheSerArgGlyIleSerAspAsnT 1087
              ::      ||||||||               |||     |||     |||      |
Db       3474 TGAGCTGCCTGGAGTCATCCTGGAAGACCATAGAACAAGAACGAAATGCCAGCTTTGTGT 3533

Qy       1087 yrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerPro---LysAlaLysG 1106
              |||||   ||||||:::   ||||||   :::      |||:::      |||        |
Db       3534 ACACCAAGGCTTTGATGGCCTATGCTTTTGCTCTGGCGGGGAACCAGAATAAGAGAGATG 3593

Qy       1106 luAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpVal- 1125
              ||    |||         |||         |||   :::|||            |||
Db       3594 AAATCCTGAAGTCACTTGATGAGGAAGCTATAAAAGAAAACAACTCTATCCACTGGAAAA 3653

Qy       1126 --SerSerGluSerLysLeuSerAspSerTrp-----------GlnProArgSerLeuA 1141
                    :::|||:::   |||:::                   |||      |||    :
Db       3654 GACCTCAGAAATCCAGGAAATCTGAACACCATTTGTACAAACCCCAGGCTTCCTCTGCTG 3713

Qy       1141 spIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluG 1161
              :::::|||:::    ||||||    :::|||:::                       |||::::::
Db       3714 AAGTAGAGATGAATGCATATGTGGTCCTGGCTCGCCTC------------ACTGCCCAGC 3761

Qy       1161 lyIlePro-----------------------------IleMetArgTrpLeuSerArgG 1171
                 |||                                  :::    ||||||::::::|
Db       3762 CAGCCCCATCCCCTGAGGATCTGACTTTGTCAATGAGCACCATCATGTGGCTCACAAAGC 3821

Qy       1171 lnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaL 1191
              ||:::||||||   ||||||||:::|||||||||||||    ||||||||     ||||
Db       3822 AACAGAATTCCAATGGTGGCTTCTCCTCCACACAGGACACTGTGGTGGCCCTTGATGCTC 3881

Qy       1191 euSerGluPheAlaAlaLeu---------MetAsnThrGluArgThrAsnIleGlnValT 1208
              ||||:::::   |||:::                 :::   :::   |||      :::   :::
Db       3882 TGTCCAAATATGGAGCAGTGACTTTTCGAGAAGTCAGAAAACTACTTTGGTGACTATCC 3941

Qy       1208 hrValThrGlyProSerSerPro------------------SerProLeuAlaValValG 1222
                  ||||||     |||                             |||        ::::::|
Db       3942 AATCTACAGGGTCATTTTCCCAAAAGTTCCAAGTGGAGAACAGTAATCGCCTGTTACTAC 4001

Qy       1222 lnProMetAlaVal--------------------AsnIleSerAlaAsnGlyPheGlyP 1235
              ||    :::|||:::                        ||||||     :::|||     |||
Db       4002 AGCAGGTCGCATTACCAGACATTCCTGGGGACTATACCATCAGTGTGTCAGGGGAAGGAT 4061

Qy       1235 heAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgA 1255
                         |||       :::   ||||||:::
Db       4062 GTGTGTATGCTCAGACTATGCTGAGATACAACATGCACTTGGAG---------------- 4105

Qy       1255 rgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLys---------- 1271
                          :::            ||||||   :::   |||
Db       4106 -----------AAGCAGCTGTCTGCATTTGCTATATGGGTACAGACAGTACCTCTAACTT 4154

Qy       1272 --GluAsnLysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyP 1291
                    |||          |||         :::::::::         |||::::::|||
Db       4155 GTAATAACCCCAAAGGCCACAACAGCTTCCAGATCTCACTAGAAATCAGTTACACAGGCA 4214

Qy       1291 ro---GlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValP 1310
                        |||   |||    :::   :::|||    :::||||||||||
Db       4215 GCCGGCCAGCCTCCAACATGGTGATTGCTGATGTGAAGATGCTATCTGGTTTC------- 4267

Qy       1310 roSerGluAlaIleSerLeuSerGluThrValLysLysVal------------------- 1323
                          |||    |||        ||||||||||||:::
Db       4268 -----------ATCCCATTGAAACCAACAGTGAAAAAGCTTGAAAGATTAGAGCACGTGA 4316

Qy       1324 --------GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrG 1341
                      |||       :::       :::    :::||||||||     |||
Db       4317 GCAGAACAGAAGTGAGCAACAACAATGTCTTAATATATTTGGATCAGGTGACCAATCAGA 4376

Qy       1341 lnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaS 1361
                          :::                 ::::::      |||    |||    |||    |||
Db       4377 CACTGGCCTTCTCCTTCATCATTCAACAAGATATCCCAGTAAGGAACCTGCAGCCTGCCA 4436

Qy       1361 erValSerIleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSer 1380
                 |||    :::    ||||||||||||                 |||            |||:::|||
Db       4437 TTGTGAAAGTCTATGACTACTATGAGACAGATGAAATGGCTTTTGCTGAATACAGCAGC 4495
```

```
RESULT 7
BU517328
LOCUS       BU517328                 879 bp    mRNA    linear   EST 12-SEP-2002
DEFINITION  AGENCOURT_10164521 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6514905
            5', mRNA sequence.
ACCESSION   BU517328
VERSION     BU517328.1  GI:22824854
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14090  row: p  column: 10
            High quality sequence stop: 663.
FEATURES             Location/Qualifiers
     source          1..879
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6514905"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
BASE COUNT      248 a    171 c    206 g    253 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               1.64e-126     Length:          879
Score:                   1280.00       Matches:         263
Percent Similarity:      91.38%        Conservative:    2
Best Local Similarity:   90.69%        Mismatches:      6
Query Match:             17.42%        Indels:          19
DB:                      13            Gaps:            2

US-10-020-095-4 (1-1428) x BU517328 (1-879)

Qy       1070 GluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeu 1089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 GAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCTA 61

Qy       1090 AlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsn 1109
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 GCCCTTATAACTTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAAT 121

Qy       1110 MetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrpValSerSerGluSer 1129
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCC 181

Qy       1130 LysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeu 1149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 AAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTG 241
```

```
Qy       1150 LeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSer 1169
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 CTCTCACACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGC 301

Qy       1170 ArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLys 1189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 AGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAAG 361

Qy       1190 AlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrVal 1209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 GCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTG 421

Qy       1210 ThrGlyProSerSerProSerPro------------------------------------ 1217
              ||||||||||||||||||||||||
Db        422 ACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTTCTGATTGACACACACAACCGCTTACTC 481

Qy       1218 ---------------LeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGly 1232
                             |||||||||||||||||||||||||||||||||||||||||||||
Db        482 CTTCAGACAGCAGAGCTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGT 541

Qy       1233 PheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSer 1252
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        542 TTTGGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCT 601

Qy       1253 ArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGlu 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        602 AGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAA 661

Qy       1273 AsnLysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGly 1292
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        662 AATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGT 721

Qy       1293 ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGlu 1312
              ||||||||||||||||||||||||||||||||||||||||    ||||||||||||||||
Db        722 AGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGGCTTATGGTGCCTTCAGAA 781

Qy       1313 AlaIleSer-LeuSerGluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLe 1332
              ||||||||| ||||||||||||||||:::||||||      ||||||||||||||
Db        782 GCAATTTCTCCTGAGCGAGACAGTGAGGAAAGTGGGATATGATCATGGGAAAACTCAACC 841

Qy       1332 uTyrLeu---AspSerValAsnGluThr 1340
                    :::   ||||||||||||||||
Db        842 CTCTATTTNAGATTCTGTAAATGAAACC 869

RESULT 8
BQ028040/c
LOCUS       BQ028040                 667 bp    mRNA    linear   EST 27-MAR-2002
DEFINITION  UI-H-CO0-arh-c-08-0-UI.s1 NCI_CGAP_Sub9 Homo sapiens cDNA clone
            IMAGE:3106526 3', mRNA sequence.
ACCESSION   BQ028040
VERSION     BQ028040.1  GI:19763319
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 667)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..667
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3106526"
                     /tissue_type="mixed"
                     /dev_stage="mixed"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub9"
                     /note="Vector: pT7T3-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoR I; Site_2: Not I; tissues:
                     Cholonic mucosa with Crohns disease, Cholonic mucosa with
                     ulcerative colitis, Fetal thymus, Cervix, Cervical
                     adenosquamous carcinoma, Ligament cells, Prostate
                     carcinoma, Bladder carcinoma, Brain oligodenroga ;
                     NCI_CGAP_Sub9 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are CGTC, AACG, GGGCC,
                     GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional
                     information, contact: Bento Soares, bento-soares@uiowa.edu
                     TAG_LIB=UI-H-CO0
                     TAG_TISSUE=Cervical Adenosquamous Carcinoma
                     TAG_SEQ=CGAAG"
BASE COUNT      194 a    113 c    139 g    221 t
ORIGIN

Alignment Scores:
Pred. No.:              3.71e-109      Length:         667
Score:                  1117.00        Matches:        219
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            15.20%         Indels:         0
DB:                     12             Gaps:           0

US-10-020-095-4 (1-1428) x BQ028040 (1-667)

Qy         60 AlaGluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyVal 79
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        665 GCGGAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTC 606

Qy         80 PheGluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAsp 99
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        605 TTTGAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGAT 546

Qy        100 GluIleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSer 119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        545 GAGATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGT 486

Qy        120 ThrArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeu 139
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 ACCCGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTA 426

Qy        140 TyrLysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysPro 159
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 TACAAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCT 366

Qy        160 TyrLysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrp 179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 TACAAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGG 306

Qy        180 LeuSerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisPro 199
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        305 TTGTCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCA 246

Qy        200 IleLeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPhe 219
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        245 ATACTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTT 186
```

```
Qy        220 GlnValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCys 239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        185 CAGGTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGT 126

Qy        240 SerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysPro 259
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        125 TCTATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCA 66

Qy        260 ValLysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLys 278
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         65 GTGAAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTGGGGAAAAAAAAAA 9

RESULT 9
AL576908/c
LOCUS       AL576908                1201 bp    mRNA    linear   EST 01-JUN-2003
DEFINITION  AL576908 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 3-PRIME, mRNA sequence.
ACCESSION   AL576908
VERSION     AL576908.2  GI:31315187
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12939516.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI081DA09NP1&cluster=3578.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI081DA09NP1.
FEATURES             Location/Qualifiers
     source          1..1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      344 a    248 c    224 g    321 t     64 others
ORIGIN

Alignment Scores:
Pred. No.:              1.6e-98        Length:          1201
Score:                  1022.00        Matches:         222
Percent Similarity:     88.84%         Conservative:    1
Best Local Similarity:  88.45%         Mismatches:      27
Query Match:            13.91%         Indels:          5
DB:                     9              Gaps:            1

US-10-020-095-4 (1-1428) x AL576908 (1-1201)

Qy       1178 PheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeu 1197
              ||||||||||||   |||||||||     ||||||                |||||||||
Db       1030 TTTGCATCTACTCA-GATACCACT---KTGCTTAAARGYTCTTYTRA-ATTGCAGCCCTA 976

Qy       1198 MetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro 1217
              |||   ||||||  |||||||||||||||||||||||||||   |||||||||||||||
Db        975 ATGAAWACAGAARG-ACAAATATCCAAGTGACCGTGACGGGCYT-AGCTCACCAAGTCCT 918

Qy       1218 LeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIle 1237
              ||||||||    ||||||    ||||||||||||||||||||||||||||||||||||||
Db        917 CTCGCTGTGGAACAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTGGATTTGCTATT 858

Qy       1238 CysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArgSer 1257
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        857 TGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCT 798

Qy       1258 IleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLysAspAspLeu 1277
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        797 ATCCAAAATCAAGAAGCCTTTGATTTAGATGTCGCTGTAAAAGAAAATAAAGATGATCTC 738

Qy       1278 AsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAla 1297
              ||||||||||||||||||||   |||||||||||||||||||||||||||||||||||||
Db        737 AATCATGTGGATTTGAATGTGTGAACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCT 678

Qy       1298 LeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSer 1317
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        677 CTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGC 618

Qy       1318 GluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerVal 1337
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        617 GAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTA 558

Qy       1338 AsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThr 1357
              |||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        557 AATGAAACCCAGTTTTGTGTTAATATNCCTGCTGTGAGAAACTTTAAAGTYTCAAATACC 498

Qy       1358 GlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaValArgSer 1377
              |||:::|||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db        497 CAAGAGGCTTCAGTGTCCATAGTGGATAACTATGAGCCAAGGAGACAGGCGGTGAGAAGT 438

Qy       1378 TyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArg 1397
                 ||||||||||||   |||      |||||||||   ||||||    |||      |||
Db        437 AACAACTCTGAAGTGAMACTGCCCCCCTGTGACCTTCGCAGTGATGMCCAGVGCMSCCGM 378

Qy       1398 ProCysGluAspGlyAlaSerGlySerHisHisHisSerSerValIlePheIlePheCys 1417
              ||||||||||||||||||||||    ||||||||||||||||||||||||||||||||||
Db        377 CCTTGTGAGGATGGAGCVTCAVGCMCCCATCATCACTCRTCAGTCATTTTTATTTTCTGT 318

Qy       1418 PheLysLeuLeuTyrPheMetGluLeuTrpLeu 1428
                 |||||||||   |||||||||||||||||
Db        317 CTCAAGCTTCTGAACTTTATGGAACTTTGGCTG 285

RESULT 10
AK042169
LOCUS       AK042169                1495 bp    mRNA    linear   HTC 05-DEC-2002
DEFINITION  Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched
            library, clone:A630065N19 product:weakly similar to
            alpha-2-macroglobulin (fragments) [Limulus polyphemus], full insert
            sequence.
ACCESSION   AK042169
VERSION     AK042169.1  GI:26334978
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
```

```
  PUBMED    11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
  PUBMED    11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
  PUBMED    11217851
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 1495)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
            Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
            Trust/MRC building Addenbrookes Hospital Cambridge) whose
            assistance we gratefully acknowledge.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..1495
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:A630065N19"
                     /db_xref="taxon:10090"
                     /clone="A630065N19"
                     /tissue_type="thymus"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="3 days neonate"
     CDS             121..954
                     /note="unnamed protein product; putative
                     weakly similar to alpha-2-macroglobulin (fragments)
                     [Limulus polyphemus] (PIR|A36260, evidence: FASTY,
                     51.8%ID, 76.7%length, match=168)"
                     /codon_start=1
                     /protein_id="BAC31190.1"
                     /db_xref="GI:26334979"
                     /translation="MRWLIQQRNSLGGFVSTQDTVVALKALSEFSALVHKENTDIQLT
                     VTGPGIPRSIHFRIDSQNLFLLHQEELHALDPITVNVSAHGSGFAICQLNVDYNVKGS
                     GSSKRRRSTENQEVFDLDVIVNNEDDISHLNLNVCTSHLGSERTGMVLMEVNLLSGFS
                     ASSDSIPLSETLKKVEYDNGKLNLYLDSVNESQFCVNIPTVRDYKVSNIRDGSVSVMD
                     YYEPRRQAVRSYNTQVKLSSCYLSPDTNCKSHTDGATDSLRRSSSLLVFCSVLLYFVQ
                     H"
BASE COUNT     368 a    349 c    334 g    444 t
ORIGIN

Alignment Scores:
Pred. No.:              9.39e-97       Length:         1495
Score:                  1007.00        Matches:        202
Percent Similarity:     76.18%         Conservative:   41
Best Local Similarity:  63.32%         Mismatches:     56
Query Match:            13.70%         Indels:         20
DB:                     11             Gaps:           4

US-10-020-095-4 (1-1428) x AK042169 (1-1495)

Qy       1124 TrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGlu 1143
              |||:::||||||          ||||||   ||||||||||||||||:::|||||||||
Db          1 TGGTTGTCATCAGGACCTGCGCTCTCTGGGTCCTGGCAGCCGCGCTCCGTGGACATCGAA 60

Qy       1144 ValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIlePro 1163
              :::||||||||||||||||||:::|||               ||||||||||||||
Db         61 ATCGCTGCTTATGCACTGCTGGCGCACACTCTGCAC---CATGTGTCTGAGGGAATCCCG 117

Qy       1164 IleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAsp 1183
              :::||||||||||||   :::|||||||||||||||||||||||||   ||||||||||||
Db        118 GTTATGAGGTGGCTCATCCAGCAAAGAAACAGCCTGGGAGGTTTTGTATCCACTCAGGAT 177

Qy       1184 ThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArgThr 1203
              |||   ||||||||||||||||||||||||||:::||||||::::::   |||   |||
Db        178 ACTGTTGTGGCCCTGAAGGCATTATCTGAATTTTCAGCCCTTGTGCACAAGGAAAATACA 237

Qy       1204 AsnIleGlnValThrValThrGlyProSerSerPro------------------------ 1215
              :::||||||:::||||||||||||||||       |||
Db        238 GATATCCAACTGACCGTCACAGGGCCCGGCATCCCCAGATCCATACACTTCCGAATTGAC 297

Qy       1216 ---------------------------SerProLeuAlaValValGlnProMetAlaVal 1226
                                         |||   :::   |||:::   |||
Db        298 TCCCAGAACCTCTTTCTCCTTCACCAGGAAGAGCTTCATGCGCTAGATCCCATCACCGTT 357

Qy       1227 AsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnVal 1246
              |||:::||||||:::|||   ||||||||||||||||||||||||||   |||||||||
Db        358 AATGTTTCTGCGCATGGCTCGGGATTTGCTATTTGCCAGCTTAATGTTGACTATAACGTG 417

Qy       1247 LysAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeu 1266
              |||   |||||||||||:::|||||||||||   :::|||||||||   |||||||||
Db        418 AAAGGTTCAGGTTCTTCTAAAAGACGGAGATCTACCGAAAATCAAGAAGTTTTTGATTTA 477
```

```
Qy       1267 AspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsnValCysThr 1286
              ||||||   |||       |||:::|||||:::::|||:::::||||||||||||||
Db        478 GACGTCATTGTG---AATAATGAGGACGACATTAGTCACCTGAATCTGAATGTGTGCACA 534

Qy       1287 SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly 1306
              |||       |||       |||:::||||||   |||||||||||||||||||||||||
Db        535 AGTCACTTGGGTTCAGAGAGGACAGGCATGGTCCTTATGGAAGTGAACCTTCTCAGTGGC 594

Qy       1307 PheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysValGluTyrAsp 1326
              |||       |||::::::|||   ||||||||||||:::||||||||||||||||
Db        595 TTTAGTGCATCTTCAGATTCAATTCCTCTGAGTGAGACCCTGAAGAAAGTGGAATATGAT 654

Qy       1327 HisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIle 1346
              :::||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db        655 AACGGGAAACTCAACCTTTATTTAGATTCTGTAAATGAATCCCAGTTTTGTGTTAATATT 714

Qy       1347 ProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAsp 1366
              |||   ||||||::::::||||||||||||   :::|||   ||||||||::::::|||
Db        715 CCCACTGTGAGAGACTACAAAGTTTCAAATATTCGAGATGGTTCAGTATCTGTAATGGAC 774

Qy       1367 TyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSer 1386
              |||||||||||||||||||||||||||||||||||||||||:::::::||||||||||||
Db        775 TACTATGAACCAAGGAGGCAGGCAGTGCGAAGCTACAACACCCAGGTGAAGCTGTCCTCA 834

Qy       1387 CysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAlaSerGlySer 1406
              |||   |||       |||       |||:::         |||||||||:::   |||
Db        835 TGTTACCTCAGTCCAGACACC---AACTGCAAGTCACACACGGACGGAGCCACGGACTCC 891

Qy       1407 HisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPheMetGlu 1425
                          ||||||   :::   :::||||||      ||||||||||||::::::
Db        892 CTTCGACGTTCTTCGAGCCTTCTCGTCTTTTGCTCCGTCCTTCTGTACTTTGTGCAA 948
```

RESULT 11
BX386457

```
LOCUS       BX386457                 978 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  BX386457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 5-PRIME, mRNA sequence.
ACCESSION   BX386457
VERSION     BX386457.1  GI:30457430
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 978)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS1AI021ZB10QP1&cluster=3578.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS1AI021ZB10QP1.
FEATURES             Location/Qualifiers
     source          1..978
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      374 a    151 c    179 g    220 t     54 others
ORIGIN
```

```
Alignment Scores:
Pred. No.:               6.09e-97      Length:         978
Score:                   1006.00       Matches:        217
Percent Similarity:      75.59%        Conservative:   9
Best Local Similarity:   72.58%        Mismatches:     72
Query Match:             13.69%        Indels:         2
DB:                      13            Gaps:           0
```

US-10-020-095-4 (1-1428) x BX386457 (1-978)

```
Qy          3 GlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeuAlaVal 22
              ||||||   ||||||      |||   :::||||||||||||||||||||||||||||||
Db         84 GGGCCANCGCTCCTGAACGNAGCACAAATCCTCTGCGTGTGCACCGCCGCGCTGGCCGTG 143

Qy         23 AlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGlyAsnVal 42
              |||||||||   ||||||   |||   |||||||||||||||||||   |||||||||||
Db        144 GCTCCMGGGCMTCGGTTTMTGGTGAMAGCMCCAGGGATAATMAGGCACGGAGGAAATGTG 203

Qy         43 ThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValLysAlaGluLeu 62
              |||||||||||||||||||||||||||||   |||||||||||||||||||||
Db        204 ACWATTGGGGTGGAGCTTCTGGAACACTGCMCTTCACAGGTGACTGTGAAGGCGGA-SWG 262

Qy         63 LeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPheGluLys 82
              |||||||||||            |||||||||:::|||   |||   |||||||||||
Db        263 CTCAAGACAGCATAAAAAMTMAAAGTATCAGTAATGGAAGAAGAAAGAGTMTTTGAAAAA 322

Qy         83 GlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyr 102
                 |||   |||                           ||||||||||||||
Db        323 AGATCTWTTAARAAAAWAAAWAATCAAAAAAAAAAWAWAAACAGTGCAGATGARAWTAAT 382

Qy        103 GluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeu 122
              :::      |||   |||   |||||||||||         |||      |||
Db        383 AARCWAAGTGTAAAAGGAAGTACCCAGGATGARAWTWTAATCTCTAAAWAGAACAMGCWTA 442

Qy        123 SerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysPro 142
                    |||||||||:::|||             :::      |||   |||   ||||||
Db        443 AAAATTGAAACCAAAAAAATAAAAGAAWTAAAAAAAAAARAAAAGGAATTAAACAAGCCA 502

Qy        143 LysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThr 162
              ||||||||||||||||||||||||||||||||||||||            |||   ||||||
Db        503 AAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTAARAATTAAAGCATTACAAAAAA 562

Qy        163 SerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGln 182
              :::||||||||:::||||||:::      |||   |||||||||||||||||||||||||||
Db        563 ACTTTAAACATTATCATTAAGGAAACCAAATAAAATTTGATCCAACAGTGGTTGTCACAA 622

Qy        183 GlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGly 202
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 CAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATACTTGGT 682

Qy        203 AspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnValSer 222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 GACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAGGTTTCA 742

Qy        223 GluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsn 242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 GAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCTATGAAT 802

Qy        243 SerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGly 262
              |||||||||||||||||||||||   ||||||                |||      |||
Db        803 TCTAAGCATTTAAATGGTACYATCATTGCAAARSATATAGATTSTAAGTYAGYGAAATCA 862

Qy        263 AspValThrLeuThrPheLeuProLeuSerPheTrpGlyLys-LysLysAsnIleThrLy 282
              |||:::||||||||||||   |||         |||||||||     ||||||||||||||
Db        863 GACTTAACTCTKACATTTTKACCTTTWKCCKTGTGGGGSAAAKAAKAAAAATATTACAAA 922

Qy        282 sThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        923 AACATTTAAGATAAATGSATCTGCAAAYTTCTCTTTTAATGATGAAGAGATGAAA 977
```

```
RESULT 12
BG173799
LOCUS       BG173799                 836 bp    mRNA    linear   EST 06-FEB-2001
DEFINITION  602333853F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4457172 5',
            mRNA sequence.
ACCESSION   BG173799
VERSION     BG173799.1  GI:12680502
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 836)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10253  row: i  column: 13
            High quality sequence stop: 694.
FEATURES             Location/Qualifiers
     source          1..836
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4457172"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /clone_lib="NCI_CGAP_Mam1"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT     228 a    179 c    202 g    227 t
ORIGIN

Alignment Scores:
Pred. No.:              3.37e-94       Length:         836
Score:                  979.50         Matches:        200
Percent Similarity:     87.55%         Conservative:   25
Best Local Similarity:  77.82%         Mismatches:     30
Query Match:            13.33%         Indels:         4
DB:                     10             Gaps:           1

US-10-020-095-4 (1-1428) x BG173799 (1-836)

Qy        942 LysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTyr 961
              :::||||||||||   ||||||||||||||||||||||||:::|||||||||||||||
Db          1 CAGAAACAGCTGACAGTTAATTTAAAAGAAAAGGCCCTTTCATATATGAGGCAAGGTTAC 60

Qy        962 GlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAsp 981
              |||||||||||||||||||||||||||||||||||||||||||||||||:::   |||
Db         61 CAAAGGGAGCTTCTCTATCAGAGGGAAGATGGCTCCTTCAGTGCTTTTGGGGACATTGAC 120

Qy        982 ProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspPro 1001
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TCTTCTGGGAGCACTTGGCTGTCAGCATTTGTTTTAAGATGCTTTCTGGAAGCTGATTAC 180

Qy       1002 TyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGln 1021
              ||||||||||||||||||:::||||||||||||||||||||||||||     |||:::
Db        181 TATATAGATATTGATCAGGATGTGTTACACAGAACATATACTTGGCTTAATGCACATAAG 240

Qy       1022 LysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGly 1041
              |||   ||||||||||||||:::||||||||||||||||||||||||||||||||||||
Db        241 AAATTCAATGGTGAATTTTGGGAGCCAGGAAGAGTGATTCACAGTGAACTTCAAGGTGGC 300

Qy       1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061
                 ||||||||||||||||||||||||||||||||||||:::||||||::::::|||
Db        301 ACCAAAAGCCCAGTAACACTTACGGCCTATATTGTGACTTCTGTCCTGGGACACAAAAAG 360

Qy       1062 TyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSerArg 1081
              ||||||||||||||||||||||||:::||||||    |||    |||    |||||||||||
Db        361 TATCAGCCTAATATCGATGTACAAGACTCAATCAAGTTTTG-GAATTTGAATTCAGCAGA 419

Qy       1082 GlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer 1101
              |||||||||||||||||||||||||:::|||:::|||||||||||||:::||||||||
Db        420 GGAATTTCGGACAATTATACCTTAGCAATTATATCCTATGCCCTGTCCACAGTCGGGAGC 479

Qy       1102 ProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMet 1121
              ||||||||:::||||||||||||||||:::|||       |||:::|||:::||||||
Db        480 CCTAAAGCAGAGGAGGCTTTGAACTTGCTGATGCAGCGATCAGAAAAGGAAGGAGACACA 539

Qy       1122 GlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAsp 1141
              ||||||||:::||||||          ||||||  ||||||||||||||||||:::|||
Db        540 CAGTTCTGGTTGTCATCAGGACCTGCGCTCTCTGGGTCCTGGCAGCCGCGCTCCGTGGAC 599

Qy       1142 IleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGly 1161
              ||||||:::|||||||||||||||||::::::|||    |||            |||    |||
Db        600 ATCGAAATCGCTGCTTATGCACTGGTGGCGCACACTCTGCAC---CATGTGTCTTGAGGA 656

Qy       1162 IleProIleMetArgTrpLeu-SerArgGlnArgAsnSerLeuGlyGlyPheAlaSerTh 1181
              ||||||:::|||:::||||||  |||||||||||||||||||||||||||    |||    |||||
Db        657 ATCCCGGTTATGAAGTGGCTCATCCAGGCAAAGAAACAGCCTGGGAGG-TTTGTATCCAC 715

Qy       1181 rGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeu 1197
              ||||||      |||    ||||||||    |||||||||:::||||||
Db        716 TCAGGATATTGTTGTGGGCCTGAAGGCATTTTCTGAATTTTCCGCCTTG 764

RESULT 13
BU624356/c
LOCUS       BU624356                 638 bp    mRNA    linear   EST 23-SEP-2002
DEFINITION  UI-H-FG1-bgj-l-22-0-UI.s1 NCI_CGAP_FG1 Homo sapiens cDNA clone
            UI-H-FG1-bgj-l-22-0-UI 3', mRNA sequence.
ACCESSION   BU624356
VERSION     BU624356.1  GI:23290571
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 638)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..638
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FG1-bgj-l-22-0-UI"
                     /tissue_type="Cell lines"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FG1"
                     /note="Organ: Enchondroma; Vector: pT7T3-Pac (Pharmacia)
                     with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
```

```
                NCI_CGAP_FG1 is a normalized cDNA library obtained from a
                pool of mRNA from 2 cell lines from Enchondroma tissues.
                The library was constructed according to Bonaldo, Lennon
                and Soares, Genome Research, 6:791-806, 1996. First strand
                cDNA synthesis was primed with an oligo-dT primer
                containing a Not I site.  Double stranded cDNA was ligated
                to an EcoR I adaptor, digested with Not I, and cloned
                directionally into pT7T3-Pac vector. The oligonucleotide
                used to prime the synthesis of first-strand cDNA contains
                a library tag sequence that is located between the Not I
                site and the (dT)18 tail. The sequence tag for this
                library is CGGTCACTC. The cell lines were provided by Dr.
                James Martin from the University of Iowa.
                TAG_LIB=UI-H-FG1
                TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
                TAG_SEQ=CGGTCACTC"
BASE COUNT      203 a    139 c    109 g    187 t
ORIGIN

Alignment Scores:
Pred. No.:               1.24e-93     Length:        638
Score:                   972.50       Matches:       190
Percent Similarity:      99.48%       Conservative:  1
Best Local Similarity:   98.96%       Mismatches:    1
Query Match:             13.23%       Indels:        1
DB:                      13           Gaps:          0

US-10-020-095-4 (1-1428) x BU624356 (1-638)

Qy       1237 IleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgArg 1256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        636 ATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGA 577

Qy       1257 SerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLysAspAsp 1276
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 TCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGAT 517

Qy       1277 LeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMet 1296
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        516 CTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATG 457

Qy       1297 AlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeu 1316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 GCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTG 397

Qy       1317 SerGluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSer 1336
              |||||   :::|||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 AGCGAGA--TTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCT 339

Qy       1337 ValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsn 1356
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        338 GTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAAT 279

Qy       1357 ThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaValArg 1376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        278 ACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGA 219

Qy       1377 SerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCys 1396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        218 AGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGC 159

Qy       1397 ArgProCysGluAspGlyAlaSerGlySerHisHisHisSerSerValIlePheIlePhe 1416
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        158 CGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTC 99

Qy       1417 CysPheLysLeuLeuTyrPheMetGluLeuTrpLeu 1428
              ||||||||||||||||||||||||||||||||||||
Db         98 TGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTG 63

RESULT 14
BB613975
LOCUS       BB613975                 648 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB613975 RIKEN full-length enriched, 0 day neonate head Mus
            musculus cDNA clone 4832419I01 5', mRNA sequence.
ACCESSION   BB613975
VERSION     BB613975.1  GI:16454440
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .648
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4832419I01"
                     /sex="mixed"
                     /tissue_type="head"
                     /dev_stage="0 day neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 0 day neonate
                     head"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
```

```
                 prepared by using trehalose thermo-activated reverse
                 transcriptase and subsequently enriched for full-length by
                 cap-trapper. cDNA went through one round of normalization
                 to Rot = 10.0 and subtraction to Rot = 100.0. Second
                 strand cDNA was prepared with the primer adapter of
                 sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC
                 3']. cDNA was cloned into the XhoI and BamHI sites.
                 Vector: a modified pBluescript KS(+) after bulk excision
                 from Lambda FLC I."
BASE COUNT      195 a    118 c    140 g    194 t       1 others
ORIGIN

Alignment Scores:
Pred. No.:                3.12e-90       Length:        648
Score:                    941.00         Matches:       178
Percent Similarity:       92.09%         Conservative:  20
Best Local Similarity:    82.79%         Mismatches:    17
Query Match:              12.81%         Indels:        0
DB:                       10             Gaps:          0

US-10-020-095-4 (1-1428) x BB613975 (1-648)

Qy        862 GlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSer 881
              :::|||:::|||||||||||||||||:::   ::::::|||      :::::::::
Db          2 AAATCTGTCTTATTGGATCTGACAGATAGCAACGTAGAAAGTAAGCAGCAATCTATGAGA 61

Qy        882 PheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGly 901
              ||||||||||||:::||||||   |||||||||||||||||||||||||||||||
Db         62 TTCTCCTTTCCTCCTGATACGGTCATTGGCAGTGAAAGAGTTCAGATCACAGCAATTGGA 121

Qy        902 AspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCys 921
              |||:::||||||   |||||||||||||||||||||||||||||||||||||||||
Db        122 GATATCCTTGGTTCCTCCATCAATGGCTTATCTTCACTGATCCGGATGCCTTATGGATGT 181

Qy        922 GlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLys 941
              ||||||||||||||||||   ||||||||||||||||||||||||||||||||||||||
Db        182 GGTGAACAGAACATGATATATTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAA 241

Qy        942 LysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTyr 961
              :::||||||||||||   ||||||||||||||||||||||||:::||||||||||||||
Db        242 CAGAAACAGCTGACAGTTAATTTAAAAGAAAAGGCCCTTTCATATATGAGGCAAGGTTAC 301

Qy        962 GlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAsp 981
              |||||||||||||||||||||||||||||||||||||||||||||||||:::   |||
Db        302 CAAAGGGAGCTTCTCTATCAGAGGGAAGATGGCTCCTTCAGTGCTTTTGGGGACATTGAC 361

Qy        982 ProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspPro 1001
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 TCTTCTGGGAGCACTTGGCTGTCAGCATTTGTTTTAAGATGCTTTCTGGAAGCTGATTAC 421

Qy       1002 TyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHisGln 1021
              ||||||||||||||||||:::||||||||||||||||||||||||||||      |||:::
Db        422 TATATAGATATTGATCAGGATGTGTTACACAGAACATATACTTGGCTTAATGCACATAAG 481

Qy       1022 LysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyGly 1041
              |||   |||||||||||||||:::|||||||||||||||||||||||||||||||||||
Db        482 AAATTCAATGGTGAATTTTGGGAGCCAGGAAGAGTGATTCACAGTGAACTTCAAGGTGGC 541

Qy       1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061
                 ||||||||||||||||||||||||||||||||||||:::|||||||      :::|||
Db        542 ACCAAAAGCCCAGTAACACTTACGGCCTATATTGTGACTTCTGTCCTGGGATANCAAAAG 601

Qy       1062 TyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGlu 1076
              ||||||||||||||||||||||||:::||||||   ||||||||
Db        602 TATCAGCCTAATATCGATGTACAAGACTCAATCAAGTTTTTGGAA 646

RESULT 15
AA418644
LOCUS       AA418644               634 bp    mRNA    linear   EST 16-OCT-1997
DEFINITION  zv93f07.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:767365 5'
            similar to TR:G534873 G534873 ALPHA-2-MACROGLOBULIN. ;, mRNA
            sequence.
ACCESSION   AA418644
VERSION     AA418644.1  GI:2080463
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 494.
FEATURES             Location/Qualifiers
     source          1. .634
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:767365"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      174 a    109 c    139 g    212 t
ORIGIN

Alignment Scores:
Pred. No.:                3.99e-87       Length:        634
Score:                    912.00         Matches:       174
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              12.41%         Indels:        0
DB:                       9              Gaps:          0

US-10-020-095-4 (1-1428) x AA418644 (1-634)

Qy       1255 ArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLys 1274
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAA 60

Qy       1275 AspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSer 1294
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGT 120

Qy       1295 GlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIle 1314
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATT 180

Qy       1315 SerLeuSerGluThrValLysLysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeu 1334
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTA 240
```

```
Qy       1335 AspSerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysVal 1354
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTT 300

Qy       1355 SerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAla 1374
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCG 360

Qy       1375 ValArgSerTyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGln 1394
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAG 420

Qy       1395 GlyCysArgProCysGluAspGlyAlaSerGlySerHisHisHisSerSerValIlePhe 1414
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTT 480

Qy       1415 IlePheCysPheLysLeuLeuTyrPheMetGluLeuTrpLeu 1428
              ||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTG 522
```

Search completed: January 17, 2004, 02:52:00
Job time : 5723 secs

```
OM protein - protein search, using sw model

Run on:          January 15, 2004, 18:02:49 ; Search time 54 Seconds
                                             (without alignments)
                                             4197.436 Million cell updates/sec

Title:           US-10-020-095-4
Perfect score:   7348
Sequence:        1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters:      1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_19Jun03:*
                 1:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                 2:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                 3:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                 4:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                 5:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                 6:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                 7:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                 8:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                 9:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                 10:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                 11:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                 12:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                 13:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                 14:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                 15:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                 16:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                 17:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                 18:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                 19:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                 20:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                 21:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                 22:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                 23:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
                 24:  /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7331	99.8	1428	24	AAE32012	Human r150 protein
2	7326.5	99.7	1445	23	ABB82165	Human CD109 K1 pro
3	7326.5	99.7	1445	23	ABB82167	Human CD109 K1-H7
4	7326.5	99.7	1445	23	AAO19372	Human platelet all
5	7318.5	99.6	1445	24	AAE32013	Human r150 protein
6	7317.5	99.6	1445	23	ABB82166	Human CD109 K1 var
7	7317.5	99.6	1445	23	ABB82168	Human CD109 K1-H7
8	7317.5	99.6	1445	23	AAO19373	Human platelet all
9	7307.5	99.4	1445	21	AAB12127	Hydrophobic domain
10	7081	96.4	1382	24	AAE32014	Human blood cell s
11	7068.5	96.2	1399	24	AAE32015	Human r150 protein
12	5470	74.4	1067	23	ABB82169	Human CD109 K15 pr
13	5461	74.3	1067	23	ABB82170	Human CD109 K15 va
14	2860.5	38.9	1445	24	AAE32018	Human blood cell s
15	1854.5	25.2	1508	22	AAB50676	C. elegans alpha-2
16	1850	25.2	1519	22	AAB50677	C. elegans alpha-2
17	1628.5	22.2	336	23	ABB90069	Human polypeptide
18	1592	21.7	352	22	ABB11581	Human secreted pro
19	1441	19.6	1474	21	AAY97157	Human alpha-2-macr
20	1440	19.6	1474	22	AAB50673	Human alpha-2 macr
21	1440	19.6	1474	23	ABP65218	Hypoxia-regulated
22	1440	19.6	1474	23	AAU74798	Human alpha 2 macr
23	1440	19.6	1474	23	AAU81017	Human alpha2 macro
24	1437.5	19.6	1508	22	AAG63548	A human alpha-2 ma
25	1437	19.6	1451	23	AAU81018	Human alpha2 macro
26	1430	19.5	1203	22	ABB66669	Drosophila melanog
27	1428.5	19.4	1474	12	AAR11334	Recombinant human
28	1425	19.4	1484	12	AAR11749	Human alpha-2 macr
29	1418.5	19.3	1492	23	ABG76898	Human alpha-2-macr
30	1413.5	19.2	1492	23	ABG61768	Novel alpha-2-macr
31	1406	19.1	1425	22	ABB71908	Drosophila melanog
32	1396.5	19.0	1487	22	ABG28006	Novel human diagno
33	1376.5	18.7	1479	23	ABB09519	Human pregnancy zo
34	1352.5	18.4	1475	22	ABG21224	Novel human diagno
35	1329.5	18.1	281	23	ABB99092	Alpha-1 proteinase
36	1328.5	18.1	1760	22	ABB63367	Drosophila melanog
37	1328.5	18.1	1760	23	ABG70020	Larval viability a
38	1315.5	17.9	1542	22	ABG28005	Novel human diagno
39	1299.5	17.7	1413	22	ABB62921	Drosophila melanog
40	1255	17.1	1285	21	AAB43949	Human cancer assoc
41	1173.5	16.0	248	22	AAB74689	Human protease and
42	1157	15.7	1411	24	ABU52372	Human GPCR related
43	1109.5	15.1	1436	24	ABU52373	Human GPCR related
44	1101.5	15.0	753	22	ABB67122	Drosophila melanog
45	1099.5	15.0	798	22	ABB62893	Drosophila melanog

ALIGNMENTS

```
RESULT 1
AAE32012
ID   AAE32012 standard; Protein; 1428 AA.
XX
AC   AAE32012;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 protein #1.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 703
FT                   /label= Ser, Tyr
FT                   /note= "Encoded by TNC"
FT   Cleavage-site   1404
FT                   /note= "GPI anchor cleavage site"
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
```

```
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   N-PSDB; AAD49434.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 1; Page 94-99; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human r150 protein.
XX
SQ   Sequence   1428 AA;

  Query Match             99.8%;  Score 7331;  DB 24;  Length 1428;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1425;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGXRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
                               O
Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
                                      O
Qy       1201 ERTNIQVTVTGPSSPSPLAVVQPMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQN 1260
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db       1201 ERTNIQVTVTGPSSPSPLAVVQPTAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQN 1260

Qy       1261 QEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETV 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 QEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETV 1320

Qy       1321 KKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 KKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380

Qy       1381 EVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYFMELWL 1428
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 EVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYFMELWL 1428


RESULT 2
ABB82165
ID   ABB82165 standard; Protein; 1445 AA.
XX
AC   ABB82165;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1 protein.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1.
XX
OS   Homo sapiens.
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
```

```
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   N-PSDB; ABQ79964.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 21; Fig 3a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1 sequence.
XX
SQ   Sequence   1445 AA;

  Query Match             99.7%;  Score 7326.5;  DB 23;  Length 1445;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 1427;  Conservative    1;  Mismatches    0;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200

Qy       1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
              |||||||||||||||||                 ||||||||||||||||||||||||||
Db       1201 ERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVV 1260

Qy       1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1320

Qy       1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1380

Qy       1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1440

Qy       1424 MELWL 1428
              |||||
Db       1441 MELWL 1445
```

```
RESULT 3
ABB82167
ID   ABB82167 standard; Protein; 1445 AA.
XX
AC   ABB82167;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1-H7 protein.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1-H7.
XX
OS   Homo sapiens.
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   N-PSDB; ABQ79966.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 21; Fig 3a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1-H7 sequence.
XX
SQ   Sequence   1445 AA;

  Query Match             99.7%;  Score 7326.5;  DB 23;  Length 1445;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 1427;  Conservative    1;  Mismatches    0;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200

Qy       1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
              |||||||||||||||||                 ||||||||||||||||||||||||||
Db       1201 ERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVV 1260

Qy       1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1320

Qy       1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1380

Qy       1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1440

Qy       1424 MELWL 1428
              |||||
Db       1441 MELWL 1445
```

```
RESULT 4
AAO19372
ID   AAO19372 standard; Protein; 1445 AA.
XX
AC   AAO19372;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human platelet alloantigen Gova.
XX
KW   Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;
KW   SNP; biallelic; bleeding disorder; post-transfusion purpura;
KW   post-transfusion platelet refractoriness; haemostatic; vaccine;
KW   neonatal alloimmune thrombocytopenia.
XX
OS   Homo sapiens.
XX
PN   WO200270738-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00291.
XX
PR   07-MAR-2001; 2001US-273941P.
XX
PA   (SCHU/) SCHUH A.
XX
PI   Schuh A,  Ouwehand W;
XX
DR   WPI; 2002-713460/77.
DR   N-PSDB; AAL49815.
XX
PT   New isolated oligonucleotide binding to a region of CD109 nucleic acid
PT   having a single nucleotide polymorphism that distinguishes a Gova
PT   and/or Govb allele, useful for treating blood disorders e.g. alloimmune
PT   thrombocytopenia  -
XX
PS   Disclosure; Page 29-35; 69pp; English.
XX
CC   The present invention relates to a sequence capable off binding
CC   specifically to a CD109 nucleic acid which has a single nucleotide
CC   polymorphism that distinguishes the Gova and Govb alleles. Detection of
CC   the Gov genotype is useful for detecting whether the subject has or is at
CC   risk of a blood disease, disorder or abnormal physical state, such as
CC   bleeding, or increased risk of bleeding, due to alloimmune destruction of
CC   blood platelets, e.g., post-transfusion purpura, post-transfusion
CC   platelet refractoriness or neonatal alloimmune thrombocytopenia. The
CC   nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
CC   individuals. The present sequence is the Gova encoded protein.
XX
SQ   Sequence   1445 AA;

  Query Match             99.7%;  Score 7326.5;  DB 23;  Length 1445;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 1427;  Conservative    1;  Mismatches    0;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200

Qy       1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
              ||||||||||||||||                  ||||||||||||||||||||||||||
Db       1201 ERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVV 1260

Qy       1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1320

Qy       1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1380

Qy       1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1440

Qy       1424 MELWL 1428
              |||||
Db       1441 MELWL 1445
```

```
RESULT 5
AAE32013
ID   AAE32013 standard; Protein; 1445 AA.
XX
AC   AAE32013;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 protein #2.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 703
FT                   /label= Ser, Tyr
FT                   /note= "Encoded by TNC"
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   N-PSDB; AAD49435.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 1; Page 103-109; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human r150 protein.
XX
SQ   Sequence   1445 AA;

  Query Match             99.6%;  Score 7318.5;  DB 24;  Length 1445;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1426;  Conservative    1;  Mismatches    1;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGXRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
```

```
Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200

Qy       1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
              |||||||||||||||||                 ||||||||||||||||||||||||||
Db       1201 ERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVV 1260

Qy       1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1320

Qy       1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1380

Qy       1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1440

Qy       1424 MELWL 1428
              |||||
Db       1441 MELWL 1445
```

```
RESULT 6
ABB82166
ID   ABB82166 standard; Protein; 1445 AA.
XX
AC   ABB82166;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1 variant protein.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1; variant.
XX
OS   Homo sapiens.
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   N-PSDB; ABQ79965.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 21; Fig 3b; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1 variant sequence.
XX
SQ   Sequence   1445 AA;

  Query Match             99.6%;  Score 7317.5;  DB 23;  Length 1445;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1426;  Conservative    1;  Mismatches    1;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGGSSPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200

Qy       1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
              |||||||||||||||||                 ||||||||||||||||||||||||||
Db       1201 ERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVV 1260

Qy       1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1320

Qy       1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1380

Qy       1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHSSVIFIFCFKLLYF 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHSSVIFIFCFKLLYF 1440

Qy       1424 MELWL 1428
              |||||
Db       1441 MELWL 1445
```

RESULT 7
ABB82168

```
ID   ABB82168 standard; Protein; 1445 AA.
XX
AC   ABB82168;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1-H7 variant protein.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1-H7; variant.
XX
OS   Homo sapiens.
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   N-PSDB; ABQ79967.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 21; Fig 3b; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1-H7 variant sequence.
XX
SQ   Sequence   1445 AA;

  Query Match             99.6%;  Score 7317.5;  DB 23;  Length 1445;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1426;  Conservative    1;  Mismatches    1;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
```

```
Qy      121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy      181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy      241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy      301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy      361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy      421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy      481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy      541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy      601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy      661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
            |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db      661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSW 720

Qy      721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy      781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
            ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db      781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy      841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy      901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy      961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy     1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy     1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy     1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200

Qy     1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
            |||||||||||||||||                 ||||||||||||||||||||||||||
Db     1201 ERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVV 1260

Qy     1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1320

Qy     1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1380

Qy     1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1440

Qy     1424 MELWL 1428
            |||||
Db     1441 MELWL 1445


RESULT 8
AAO19373
ID   AAO19373 standard; Protein; 1445 AA.
XX
AC   AAO19373;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human platelet alloantigen Govb.
XX
KW   Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;
KW   SNP; biallelic; bleeding disorder; post-transfusion purpura;
KW   post-transfusion platelet refractoriness; haemostatic; vaccine;
KW   neonatal alloimmune thrombocytopenia.
XX
OS   Homo sapiens.
XX
PN   WO200270738-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00291.
XX
PR   07-MAR-2001; 2001US-273941P.
XX
PA   (SCHU/) SCHUH A.
XX
PI   Schuh A,  Ouwehand W;
XX
DR   WPI; 2002-713460/77.
DR   N-PSDB; AAL49816.
XX
PT   New isolated oligonucleotide binding to a region of CD109 nucleic acid
PT   having a single nucleotide polymorphism that distinguishes a Gova
PT   and/or Govb allele, useful for treating blood disorders e.g. alloimmune
PT   thrombocytopenia  -
XX
PS   Disclosure; Page 41-47; 69pp; English.
XX
CC   The present invention relates to a sequence capable off binding
CC   specifically to a CD109 nucleic acid which has a single nucleotide
CC   polymorphism that distinguishes the Gova and Govb alleles. Detection of
CC   the Gov genotype is useful for detecting whether the subject has or is at
CC   risk of a blood disease, disorder or abnormal physical state, such as
CC   bleeding, or increased risk of bleeding, due to alloimmune destruction of
CC   blood platelets, e.g., post-transfusion purpura, post-transfusion
CC   platelet refractoriness or neonatal alloimmune thrombocytopenia. The
CC   nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
CC   individuals. The present sequence is the Govb encoded protein.
XX
SQ   Sequence   1445 AA;

  Query Match             99.6%;  Score 7317.5;  DB 23;  Length 1445;
```

```
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 1426;  Conservative    1;  Mismatches    1;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGGSSPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200

Qy       1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
              |||||||||||||||||                 ||||||||||||||||||||||||||
Db       1201 ERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVV 1260

Qy       1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1320

Qy       1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1380

Qy       1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHSSVIFIFCFKLLYF 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHSSVIFIFCFKLLYF 1440

Qy       1424 MELWL 1428
              |||||
Db       1441 MELWL 1445


RESULT 9
AAB12127
ID   AAB12127 standard; Protein; 1445 AA.
XX
AC   AAB12127;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Hydrophobic domain protein isolated from HT-1080 cells.
XX
KW   Human; secreted protein; membrane protein; hydrophobic domain;
KW   proliferation control; differentiation induction; material transport;
KW   biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
KW   immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
KW   haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
KW   autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200029448-A2.
XX
PD   25-MAY-2000.
XX
PF   17-NOV-1999;   99WO-JP06412.
XX
PR   17-NOV-1998;   98JP-0326255.
PR   22-DEC-1998;   98JP-0364315.
PR   16-MAR-1999;   99JP-0069811.
PR   27-APR-1999;   99JP-0119299.
PR   19-MAY-1999;   99JP-0138169.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Kimura T;
XX
DR   WPI; 2000-387753/33.
DR   N-PSDB; AAA60199, AAA62010.
XX
PT   Proteins comprising hydrophobic regions, such as secretory and membrane
PT   proteins, useful in research and diagnostics and having various
PT   activities e.g. immunomodulatory, antiinflammatory, chemokinetic,
```

```
PT   hemostatic, thrombolytic -
XX
PS   Claim 1; Page 225-230; 410pp; English.
XX
CC   Secretory proteins play important roles in the proliferation control, the
CC   differentiation induction, the material transport and the biophylaxis of
CC   cells. Membrane proteins have important roles as signal receptors, ion
CC   channels and transporters. The present sequence is a human protein which
CC   has at least one hydrophobic domain. This protein may be a secretory or a
CC   membrane protein. The present protein may have cytokine and cell
CC   proliferation/differentiation activity, immune stimulating or suppressing
CC   activity, haematopoiesis activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, anti-inflammatory activity and tumour
CC   inhibition activity. The present protein could therefore be used for
CC   treatment of autoimmune disease, Alzheimer's disease, Parkinson's
CC   disease, and cancer.
XX
SQ   Sequence   1445 AA;

  Query Match             99.4%;  Score 7307.5;  DB 21;  Length 1445;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 1424;  Conservative    1;  Mismatches    3;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSFPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||:|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAITQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080

Qy       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200

Qy       1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
              |||||||||||||||||                 |||||| |||||||||||||||||||
Db       1201 ERTNIQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPTAVNISANGFGFAICQLNVV 1260

Qy       1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1320

Qy       1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1380

Qy       1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYF 1440

Qy       1424 MELWL 1428
              |||||
Db       1441 MELWL 1445


RESULT 10
AAE32014
ID   AAE32014 standard; Protein; 1382 AA.
XX
AC   AAE32014;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human blood cell surface antigen, CD109 #1.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; blood cell syrface antigen; CD109.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 682
FT                   /label= Ser, Tyr
FT                   /note= "Encoded by TNC"
XX
```

```
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   N-PSDB; AAD49436.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 1; Page 112-117; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human blood cell
CC   surface antigen, CD109. CD109 is a variant of r150 sequence.
XX
SQ   Sequence   1382 AA;

  Query Match             96.4%;  Score 7081;  DB 24;  Length 1382;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1379;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         22 VAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFE 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFE 60

Qy         82 KGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYK 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYK 120

Qy        142 PKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPIL 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPIL 180

Qy        202 GDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVK 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVK 240

Qy        262 GDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPV 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPV 300

Qy        322 EILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLT 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLT 360

Qy        382 LEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSEL 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSEL 420

Qy        442 QLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVV 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVV 480

Qy        502 SRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKL 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKL 540

Qy        562 YWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGY 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGY 600

Qy        622 YLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLG 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 YLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLG 660

Qy        682 SSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVE 741
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        661 SSPHVRKHFPETWIWLDTNMGXRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVE 720

Qy        742 LQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINA 801
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        721 LQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNEINA 780

Qy        802 TGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYS 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYS 840

Qy        862 QSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGC 921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 QSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGC 900

Qy        922 GEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYD 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYD 960

Qy        982 PSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG 1041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 PSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG 1020

Qy       1042 NKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGS 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 NKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGS 1080

Qy       1102 PKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEG 1161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 PKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEG 1140

Qy       1162 IPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVV 1221
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 IPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVV 1200

Qy       1222 QPMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVD 1281
              || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 QPTAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVD 1260

Qy       1282 LNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQ 1341
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 LNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQ 1320

Qy       1342 FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCED 1401
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCED 1380

Qy       1402 GA 1403
              ||
Db       1381 GA 1382


RESULT 11
AAE32015
ID   AAE32015 standard; Protein; 1399 AA.
XX
AC   AAE32015;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 protein #3.
XX
```

```
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 682
FT                   /label= Ser, Tyr
FT                   /note= "Encoded by TNC"
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   N-PSDB; AAD49437.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 1; Page 120-126; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human r150 protein.
XX
SQ   Sequence   1399 AA;

  Query Match             96.2%;  Score 7068.5;  DB 24;  Length 1399;
  Best Local Similarity   98.6%;  Pred. No. 0;
  Matches 1380;  Conservative    1;  Mismatches    1;  Indels   17;  Gaps    1;

Qy         22 VAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFE 81
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 VAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFE 60

Qy         82 KGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYK 141
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYK 120

Qy        142 PKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPIL 201
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPIL 180

Qy        202 GDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVK 261
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVK 240

Qy        262 GDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPV 321
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPV 300

Qy        322 EILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLT 381
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 EILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLT 360

Qy        382 LEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSEL 441
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 LEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSEL 420

Qy        442 QLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVV 501
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 QLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVV 480

Qy        502 SRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKL 561
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKL 540

Qy        562 YWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGY 621
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 YWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGY 600

Qy        622 YLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLG 681
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 YLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLG 660

Qy        682 SSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVE 741
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        661 SSPHVRKHFPETWIWLDTNMGXRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVE 720

Qy        742 LQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINA 801
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||
Db        721 LQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNEINA 780

Qy        802 TGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYS 861
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYS 840

Qy        862 QSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGC 921
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 QSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGC 900

Qy        922 GEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYD 981
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYD 960

Qy        982 PSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG 1041
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 PSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG 1020

Qy       1042 NKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGS 1101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 NKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGS 1080

Qy       1102 PKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEG 1161
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 PKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEG 1140

Qy       1162 IPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSP---- 1217
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 IPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPVKFL 1200

Qy       1218 -------------LAVVQPMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAF 1264
                           |||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 IDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAF 1260

Qy       1265 DLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVKKVE 1324
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 DLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVKKVE 1320

Qy       1325 YDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKL 1384
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 YDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKL 1380

Qy       1385 SSCDLCSDVQGCRPCEDGA 1403
              |||||||||||||||||||
Db       1381 SSCDLCSDVQGCRPCEDGA 1399
```

```
RESULT 12
ABB82169
ID   ABB82169 standard; Protein; 1067 AA.
XX
AC   ABB82169;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K15 protein.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K15.
XX
OS   Homo sapiens.
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   N-PSDB; ABQ79968.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 21; Fig 5a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K15 sequence.
XX
SQ   Sequence   1067 AA;

  Query Match             74.4%;  Score 5470;  DB 23;  Length 1067;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1062;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQ 1063
              |||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQ 1063


RESULT 13
ABB82170
ID   ABB82170 standard; Protein; 1067 AA.
XX
AC   ABB82170;
XX
```

```
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K15 variant protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K15; variant.
XX
OS   Homo sapiens.
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   N-PSDB; ABQ79969.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 21; Fig 5b; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K15 variant sequence.
XX
SQ   Sequence   1067 AA;

  Query Match             74.3%;  Score 5461;  DB 23;  Length 1067;
  Best Local Similarity   99.8%;  Pred. No. 0;
  Matches 1061;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660

Qy        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSW 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        781 KVIIEKSDKFDILMTSNEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQ 1063
              |||||||||||||||||||||||||||||||||||||||||||
Db       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQ 1063


RESULT 14
AAE32018
ID   AAE32018 standard; Protein; 1445 AA.
XX
AC   AAE32018;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human blood cell surface antigen, CD109 #2.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
```

```
KW   therapy; blood cell syrface antigen; CD109.
XX
OS   Homo sapiens.
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Disclosure; Fig 17; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human blood cell
CC   surface antigen, CD109. CD109 is a variant of r150 sequence.
CC   Note: This sequence is stated to be encoded by human CD109 DNA shown
CC   in figure 17 of the specification, but this does not appear to be the
CC   case.
XX
SQ   Sequence   1445 AA;

  Query Match             38.9%;  Score 2860.5;  DB 24;  Length 1445;
  Best Local Similarity   59.9%;  Pred. No. 3.3e-178;
  Matches  861;  Conservative    1;  Mismatches  559;  Indels   17;  Gaps    1;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
Db          1 MGGPPLLTAAHLLCVCTAALAVAPGPAPLVTAPGIIAPGGAVTIGVGLLGHCPSGVTVLA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
Db         61 GLLLTASALTVSVLGAGGVPGLGSPLTLTLPSLPLASAAGITGLAVTGATGAGILPSAST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
Db        121 ALSPGTLAISVPIGTALALTLPLGGVLPAIVTLPSAPLPTLTSLAILILAPLSALIGGTL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
Db        181 SGGSALGVISLTPGLSSHPILGATSIGVGVAAGTTTGSPGVSGTVLPLPGVTLGTPLTCS 240

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
Db        241 MASLHLAGTITALTTTGLPVLGAVTLTPLPLSPTGLLLAITLTPLIAGSAAPSPAAGGML 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
Db        301 AVMASSAGLSGTLALSSPGPVGILTTVTGSVTGISAAVSTAVPPLGHATIIGPPATTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
Db        361 LPSLAPTATVLVTAAAGAGLTLGGAAAAVVITVTGAATTGTTSGSASGAGLMGAVGLIAT 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
Db        421 TVPGSGTPLIGPPILGASSGLGLLATPLGSLSSMAVHSLPLSPSLTTIGLLTAAGAILVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
Db        481 SPPGLVVSGALALLGLSTMVVSAGGLVAVGLGASTMPSLTPGASTTPLACVIVTTIGAAG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
Db        541 GIISAVLLIPVGLVPLALILLTTSLVLAGPSGLVSLAISVTGPASIVGIVAVALSVALMA 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEY 660
Db        601 ASAAITMGAVVHGLGLTATGTTLGMPMASPAVPGGCGLTVLTAAALTLATIAGVTAAAGT 660

Qy        661 AERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSW 720
Db        661 AGAPMGGAGGHIVAIHAPSLGSSPHVALHPPGTTITLATAMGTAITGGPGVTVPASITST 720

Qy        721 VATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEV 780
Db        721 VATGPVISGALGLGLTTTPVGLGAPGPPPIPLALPTSVIAGGGPALGITIPATLLAATGV 780

Qy        781 KVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPT 840
Db        781 LVIIGLSALPAILMTSAGIAATGHGGTLLVPSGAGATVLPPIAPTHLGGIPITVTALSPT 840

Qy        841 ASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAI 900
Db        841 ASAAVTGMILVLAGGIGLSTSGSILLALTAAALGSTLLTLSPSPPPATVTGSGAVGITAI 900

Qy        901 GDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQG 960
Db        901 GAVLGPSIAGLASLIAMPTGCGGGAMIAPAPAITILATLTLLLGLTAALLGLALSPMAGG 960

Qy        961 YQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGH 1020
Db        961 TGAGLLTGAGAGSPSAPGATAPSGSTTLSAPVLACPLGAAPTIAIAGAVLHATTTTLLGH 1020

Qy       1021 QKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFS 1080
Db       1021 GLSAGGPTAPGAVIHSGLGGGALSPVTLTATIVTSLLGTALTGPAIAVGGSIHPLGSGPS 1080

Qy       1081 RGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSL 1140
Db       1081 AGISAATTLALITTALSSVGSPLALGALAMLTTAAGGGGGMGPTVSSGSLLSASTGPASL 1140

Qy       1141 DIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNT 1200
Db       1141 AIGVAATALLSHPLGPGTSGGIPIMATLSAGAASLGGPASTGATTVALLALSGPAALMAT 1200

Qy       1201 ERTNIQVTVTGPSSPSP-----------------LAVVQPMAVNISANGFGFAICQLNVV 1243
Db       1201 GATAIGVTVTGPSSPSPVLPLIATHAALLLGTAGLAVVGPMAVAISAAGPGPAICGLAVV 1260

Qy       1244 YNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
Db       1261 TAVLASGSSAAAASIGAGGAPALAVAVLGALAALAHVALAVCTSPSGPGASGMALMGVAL 1320

Qy       1304 LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVS 1363
Db       1321 LSGPMVPSGAISLSGTVLLVGTAHGLLALTLASVAGTGPCVAIPAVAAPLVSATGAASVS 1380

Qy       1364 IVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHSSVIFIFCFKLL 1421
Db       1381 IVATTGPAAGAVASTASGVLLSSCALCSAVGGCAPCGAGASGSHHSSVIPIPCPLLL 1438

RESULT 15
AAB50676
ID   AAB50676 standard; Protein; 1508 AA.
XX
```

```
AC   AAB50676;
XX
DT   19-MAR-2001  (first entry)
XX
DE   C. elegans alpha-2-macroglobulin cDNA sequence SEQ ID NO:70.
XX
KW   Human; Caenorhabditis elegans; UNC-5; splice variant; nematode worm;
KW   protein-protein interaction; identification.
XX
OS   Caenorhabditis elegans.
XX
PN   WO200073328-A2.
XX
PD   07-DEC-2000.
XX
PF   02-JUN-2000; 2000WO-EP05108.
XX
PR   01-JUN-1999;   99GB-0012755.
XX
PA   (DEVG-) DEVGEN NV.
XX
PI   Van Criekinge W,  Roelens I,  Bogaert T,  Verwaerde P;
XX
DR   WPI; 2001-016508/02.
XX
PT   Three variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) and
PT   a human unc-5HS1 cDNA, useful in yeast two hybrid experiments for
PT   identifying unknown human cDNAs which encode proteins that interact
PT   with the human unc-5C protein -
XX
PS   Example 4; Page 197-202; 246pp; English.
XX
CC   The present invention describes 3 variants of human unc-5C cDNAs
CC   (unc-5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced
CC   unc-5C transcripts, and a human unc-5HS1 cDNA which shares homology with
CC   the Rattus norvegicus unc-5HS1 cDNA. Also described are assays based on
CC   protein-protein-interactions between the unc-5 protein and a variety of
CC   different interacting proteins. The unc-5C variant cDNAs and unc-5HS1
CC   cDNA are useful in methods for identifying compounds which reduce or
CC   inhibit the lethal phenotype associated with the expression of the
CC   unc-5 death domain in yeast. They are also useful in yeast two hybrid
CC   experiments for identifying unknown human cDNAs which encode proteins
CC   that interact with the human unc-5C protein. AAC90914 to AAC90971 and
CC   AAB50646 to AAB50693 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence   1508 AA;

  Query Match            25.2%;  Score 1854.5;  DB 22;  Length 1508;
  Best Local Similarity  31.8%;  Pred. No. 3.3e-112;
  Matches  476;  Conservative  292;  Mismatches  562;  Indels  169;  Gaps   44;

Qy         15 VCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVL 74
Db         28 VSTTAAPVKPA-TYMLVAPAVVRPDQPFSVCMNLLKQATDEDMIVRIEVRTERNETIAAR 86

Qy         75 EAEGVFEKGSFKTLTLPSLPLNSAD--EIYELRVTGRTQD-EILFSNSTRLSFETKRISV 131
Db         87 VISNL-KPGIAQTVSLSEMPAQSLTPRQSYKLYIRGETLNAELIFENENELKYDQKALSV 145

Qy        132 FIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISK 191
Db        146 FIQTDRAIYRPASLVRYRAIVVKSDLKPYVGNATIKIFDPSRNLISQTIGVTLDRGVYSG 205

Qy        192 TFQLSSHPILGDWSIQVQ----VNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLN 247
Db        206 ELQLAEETLLGDWFIEVETSNGVQDKS---SFTVDTYVLPKFEVNIKTSSFITIND-DLS 261

Qy        248 GTITAKYTYGKPVKGDVTLTF-LPLSFW-----------GKKKN----ITKTFKIN--GS 289
Db        262 VFVDAKYTYGKGVAGKAKVSLELPWHRWHAMVPTIIDENGVKKEEELMVERTVKLNRQGE 321

Qy        290 ANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDY 349
Db        322 AAVVFSNDELKR--------HKLLHEWGGGSIRIVASVTEDITEIERNATHQISTFREEV 373

Qy        350 IIEFFDYTTVLKPSLNFTATVKVTRADGN--QLTLEERRNNVVITVTQRNYTEYWSGSNS 407
Db        374 KLDVEKQGDTFKPGLTYNVVVALKQMDDTPVKATLPKR----VQVSTFYNYP--YNHDTS 427

Qy        408 GNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK------ 461
Db        428 SLQEEKETKIVEVDAHGTSVLTLQPPINCTSARIEAH-YDIGGKDNFTATPIYSSLYVEA 486

Qy        462 --SPSKTYIQLKTRDEN-IKVGSPFELVVSGNKRLKELSYMVVSRGQLV-----AVGKQN 513
Db        487 AVSPTKSFLQLLADNEGAVDVGKSLSFSLKATQPLSTITYQVMSRSNIVVSQQMTVNSEH 546

Qy        514 STMFSLTPENSWTPKACVIVY-YIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSE 572
Db        547 ATI-SFPATANMAPKSRLIVYAIIESSQEVLVDALDFKVEGIFQNQVALSIDKQAVEPGQ 605

Qy        573 KVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMF------ 626
Db        606 NVKFKVT-SDKNSFVGLLVVDQSVLLLKTGNDITREKVEQDLENYDSNNVGGGFGGPRPW 664

Qy        627 ------------------MNSFAVFQECGLWVLTDANLTKD---------YIDGVYDNA 658
Db        665 EAIDRKKRSIWRPWWGIGGSDAQSIFSNAGLVVLTDALLYREPQREFMSVMMMDGAPGMA 724

Qy        659 EYAERFMEENEGHIVDIHDFSLGSS---PHVRKHFPETWIWLDTNMGYRIYQEFEVTVPD 715
Db        725 EAA--FAAPPMG----------GSSPPPPTVRKFFPHTWIWSDLN-STSGEVEMEIEAPD 771

Qy        716 SITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLK 775
Db        772 TITSWVASTFAINEENGLGVAPTTSKLRVFRPFFIQLNLPYAVRRGEKFALLVLVFNYME 831

Qy        776 DATEVKVII--EKSDKFDILMTSSEI---NATGHQQTLLVPSEDGAT---VLFPIRPTHL 827
Db        832 KEQDVTVTLKYDKDSGYDLLKKDGTVVRRDEVGQQNVRIVSVAGGGTSKAVYFPIVPSSI 891

Qy        828 GEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPN 887
Db        892 GEIPVHISAIASQGGDAVEMNLRVDPQGYKVDRNIPFVIDLNNNSSDFS-KNLELIWPND 950

Qy        888 TVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTD 947
Db        951 VVDGSQKARLDVIGDMMGPVLNNAHKLVQMPYGCGEQNMLNLVPNILVVKYLRATNRNES 1010

Qy        948 NLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQ 1007
Db       1011 QLETKAIKFIEQGIQRELTYKRADNSFSAFGDSDKAGSTWLTAFVVRSFHHAKQYAFVDP 1070

Qy       1008 NVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG-NKSPVTLTAYIVTSLLGYRKYQPNI 1066
Db       1071 NVISRAVAFLNSQQMESGAFAERGEVHHKDMQGGAQDGGVALTAFVLISIL-----ENGM 1125

Qy       1067 DVQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWR--AEQEGGMQFW 1124
Db       1126 ENGKAVTYLEKHLDEVSGNAYTMAVVAYALQLAKSKQAGKAFENLKKHKIVEKSGDVKFA 1185

Qy       1125 VSSE--SKLSDS----WQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGF 1178
Db       1186 SAQKKVEKLKESRAYMFQARPVDIETTSYAVLSYLAQNQTSESLSIIRWLVSQRNELGGF 1245

Qy       1179 ASTQDTTVALKALSEFAALMNTERTNIQVTV-TGPSSPS------PLAVVQPM------- 1224
Db       1246 TSTQDTVMALQALSSYAAVTYSDKHTSQVTILNGKHTHSFDINIRNAIVLQSYQLSSLND 1305

Qy       1225 AVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNV 1284
Db       1306 AVSINANGTGVVFAQLSYSY---------YRDSLNDDAPFFCSQEIKEIRAG-NRLQLDL 1355

Qy       1285 CTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVK-----KVEYDHGKLNLYLDSVNE 1339
```

```
Db         1356 CCNYTRPGKSNMALAEIDALSGYRFDAEQVHTLTSIEDLQRVEMEKDDTKMNVYFNPLGG 1415

Qy         1340 TQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRP 1398
                   |::: :   ::|:: : |:  :||||:|  |   :| ::   |  : |   : | |
Db         1416 RPVCLSLYSDVTYQVADQKPANFRLVDYYDPEEQLKMTYAAKQTRSLQEKCG--EDCWP 1472
```

Search completed: January 15, 2004, 18:12:20
Job time : 59 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:11:15 ; Search time 22 Seconds (without alignments) 2746.359 Million cell updates/sec

Title: US-10-020-095-4
Perfect score: 7348
Sequence: 1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1441	19.6	1474	4	US-09-241-606-2	Sequence 2, Appli
2	973.5	13.2	1651	1	US-08-447-411-2	Sequence 2, Appli
3	931	12.7	1663	2	US-08-793-126-1	Sequence 1, Appli
4	931	12.7	1663	3	US-09-132-271-1	Sequence 1, Appli
5	931	12.7	1663	3	US-09-142-334-22	Sequence 22, Appl
6	921	12.5	1676	4	US-08-487-283A-2	Sequence 2, Appli
7	894	12.2	1642	2	US-08-662-227-2	Sequence 2, Appli
8	894	12.2	1642	4	US-09-017-947-2	Sequence 2, Appli
9	885	12.0	1642	1	US-08-447-411-45	Sequence 45, Appl
10	880	12.0	1648	2	US-08-662-227-35	Sequence 35, Appl
11	880	12.0	1648	4	US-09-017-947-35	Sequence 35, Appl
12	768.5	10.5	1333	1	US-08-447-411-76	Sequence 76, Appl
13	758.5	10.3	1333	2	US-08-662-227-34	Sequence 34, Appl
14	758.5	10.3	1333	4	US-09-017-947-34	Sequence 34, Appl
15	329	4.5	645	1	US-07-779-172A-3	Sequence 3, Appli
16	234	3.2	113	4	US-09-311-352B-2	Sequence 2, Appli
17	215.5	2.9	250	4	US-09-241-606-4	Sequence 4, Appli
18	185	2.5	183	4	US-09-311-352B-4	Sequence 4, Appli
19	180	2.4	1341	4	US-09-252-991A-26785	Sequence 26785, A
20	174.5	2.4	2366	1	US-08-480-604A-10	Sequence 10, Appl
21	174.5	2.4	2366	2	US-08-405-496A-10	Sequence 10, Appl
22	174.5	2.4	2366	3	US-08-915-136-10	Sequence 10, Appl
23	174.5	2.4	2366	4	US-08-957-310-10	Sequence 10, Appl
24	174.5	2.4	2366	4	US-10-011-366-10	Sequence 10, Appl
25	171	2.3	3290	4	US-09-328-352-5486	Sequence 5486, Ap
26	161	2.2	2199	4	US-08-793-273C-2	Sequence 2, Appli
27	161	2.2	2199	5	PCT-US95-11684-2	Sequence 2, Appli
28	160.5	2.2	403	1	US-08-118-674-1	Sequence 1, Appli
29	156	2.1	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
30	153.5	2.1	3169	4	US-09-453-702B-257	Sequence 257, App
31	152.5	2.1	1095	4	US-09-206-942-69	Sequence 69, Appl
32	152.5	2.1	1536	1	US-08-038-682-2	Sequence 2, Appli
33	152.5	2.1	1536	1	US-08-302-832-2	Sequence 2, Appli
34	152.5	2.1	1536	2	US-08-530-198-2	Sequence 2, Appli
35	152.5	2.1	1536	2	US-08-469-880-2	Sequence 2, Appli
36	152.5	2.1	1536	2	US-08-728-470-2	Sequence 2, Appli
37	152.5	2.1	1536	2	US-08-617-697-2	Sequence 2, Appli
38	152.5	2.1	1536	3	US-08-719-641-2	Sequence 2, Appli
39	152.5	2.1	1536	4	US-09-206-942-67	Sequence 67, Appl
40	152.5	2.1	1844	4	US-08-851-567B-53	Sequence 53, Appl
41	152.5	2.1	2504	4	US-08-851-567B-12	Sequence 12, Appl
42	150.5	2.0	2516	4	US-08-851-567B-47	Sequence 47, Appl
43	148	2.0	2071	3	US-09-415-522-6	Sequence 6, Appli
44	148	2.0	4563	4	US-09-108-006C-1	Sequence 1, Appli
45	146.5	2.0	1391	4	US-09-106-568E-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-241-606-2
; Sequence 2, Application US/09241606
; Patent No. 6472140
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Kovacs, Dora
; APPLICANT: Saunders, Aleister J.
; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 0609.4460003
; CURRENT APPLICATION NUMBER: US/09/241,606
; CURRENT FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-241-606-2

Query Match 19.6%; Score 1441; DB 4; Length 1474;
Best Local Similarity 29.1%; Pred. No. 5.9e-111;
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

```
Qy     12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
          || :      :|:  |:::|  | ::          |  ||  :    ||| |  |      | ::
Db     15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy     72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
           : ||| |    :|      ::| :|::|       ::| | ||:    |    |  :   :
Db     74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy    127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
           :   ||:|||::||| | ||||:|:: :| |     : :: |:||  | | | | |   :
Db    125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy    186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
          |:   :| |||  |    | : : ||        :| :   |  | |:||||||| :   |    ::
Db    185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy    244 KHLNGTITAKYTYGKPVKGDVTLTFLPL--------------------------SFWGK 276
          : :| ::    |||||| | ||::                                 |: :
Db    244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303

Qy    277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
            | || |::              :  |||        |       :   |  |  ||       :| ::
Db    304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy    332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
```

```
Db    348 TKLS-FVKVDSHFRQG---IPPFG--------------QVRLVDGKGVPIP----NKVI 384
Qy    392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
Db    385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441
Qy    439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
Db    442 EEAHHTAY-----------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489
Qy    493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
Db    490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549
Qy    541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
Db    550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608
Qy    601 ASNDITMENVVHEL-ELYNTGY-------------YLGMFMNS-------------FA 631
Db    609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 668
Qy    632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
Db    669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725
Qy    681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
Db    726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782
Qy    738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
Db    783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841
Qy    798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
Db    842 EQAPHCICANGRQ----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890
Qy    842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
Db    891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947
Qy    897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
Db    948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEIKSKAIGY 1007
Qy    957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
Db   1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067
Qy   1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
Db   1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127
Qy   1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
Db   1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186
Qy   1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
Db   1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245
Qy   1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
Db   1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304
Qy   1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
Db   1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353
Qy   1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
Db   1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413
Qy   1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
Db   1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465
```

RESULT 2
US-08-447-411-2
; Sequence 2, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1651 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-447-411-2

Query Match 13.2%; Score 973.5; DB 1; Length 1651;
Best Local Similarity 23.5%; Pred. No. 9.8e-72;
Matches 391; Conservative 278; Mismatches 649; Indels 343; Gaps 61;

```
Qy      1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELL-EHCPSQVTVK 59
Db      1 MEGMALYLVAALLIGFPGS---SHGALYTLITPAVLRTDTEEQILVEAHGDSTPKSLDIF 57
Qy     60 A----ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEIL 115
Db     58 VHDFPRKQKTLFQSRVDMNQAGSMFVT---PTIKVPAKELNKDSKQNQYVVVKVTGPQVA 114
Qy    116 FSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNL 175
Db    115 LEKVVLLSYQSG--FVFIQTDKGIYTPGSPVRYRVFSV--DHNMHRMDKTVIVE------ 164
Qy    176 IQQWLSQQSDLGVISKTFQLSSHPI----------------LGDWSIQVQVN---DQTYY 216
Db    165 ------FQTPEGIV-----VSSKPVNPSGSIRPYNLPELVSFGTWKAVAKYEHSPEESYT 213
```

```
Qy        217 QSFQVSEYVLPKFEVTLQTP---LYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSF 273
Db        214 AYFDVREYVLPSFEVRLQPSDKFLYIDGN-KNFHVSITARYLYGKKVEG-VAFVVFGVKI 271

Qy        274 WGKKKNI----TKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTE 329
Db        272 DDAKKSIPDSLTRIPIIDGDGEATLKRDTLRSRFQDLNQLVGHTLYVS------VTVITE 325

Qy        330 SVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNV 389
Db        326 SGSDMVVTEQGGIHIVTSPYQIYFTKTPKYFKPGMPYELTVYVTNPDGSPAA------H 378

Qy        390 VITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVP---QSGTFKIEFPILEDSSELQLKAY 446
Db        379 VPVVSEAIHSE-------GTTLSDGTAKLILNTPLNIQS------LPITVRTNHGDLPRE 425

Qy        447 FLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS--PFELVVSGN----KRLKELSYMV 500
Db        426 RQAIKSMTATAYQTQGGSENYLHVAITSTEIKPGDNLPVNFNVRGNANSLNQIKYFTYLI 485

Qy        501 VSRGQLVAVGK------QNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQ-- 552
Db        486 LNKGKIFKVGRQPRRDGQNLVTMNLHITPDLIPSFRFVAYYQVGNNEIVADSVWVDVKDT 545

Qy        553 ----LVFK------NKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNAS 602
Db        546 CMGTLVVKGASSRDDRIQ--------KPGAAMKIKLE-GDPGARVGLVAVDKAVYVLNDK 596

Qy        603 NDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLT--------------- 647
Db        597 YKISQAKIWDTIEKSDFGCTAGSGQNNLGVFEDAGLALTTSTNLNTKQRSAAKCPQPANR 656

Qy        648 -------------------------KDYIDGVYDN-----AEYAERFMEENEG------- 670
Db        657 RRRSSVLLLDSKASKAAQFQDQGLRKCCEDGMHENPMGYTCEKRAKYIQEGDACKAAFLE 716

Qy        671 --HIV------------------DIHDFSLGSSPHV-RKHFPETWIWL-------DTNMG 702
Db        717 CCHYIKGIRDENQRESELFLARSDFEDELFGDDNIISRSDFPESWLWLTEELTGEPNNQG 776

Qy        703 YRIYQEFEVTVP----DSITSWVATGFVISEDLGLGLTTT-------PVELQAFQPFFIF 751
Db        777 IS-----SKTVPFYLRDSITTW--------ELLAVGLSPTKGICVAEPYEITVMKDFFID 823

Qy        752 LNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVP 811
Db        824 LRLPYSVVKNEQVEIRAILYNYADEDIYVRVELIYNPAF----CSASTEGQRYRQQFPIK 879

Qy        812 SEDGATVLFPIRPTHLG--EIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLT 869
Db        880 ALSSRAVPFVIVPLEQGLHDVEVIASVRGELASDGVRKKLKVVPEGERKNIVTIIELDPS 939

Qy        870 DNRLQST--LKTLSFSFPPNTVTGSERVQITAIGD----VLGPSING--LASLIRMPYGC 921
Db        940 VKGVGGTQELTVIANKLDDKVPDTEVETRISVLGDPVAQIIENSIDGSKLNHLIITPSGC 999

Qy        922 GEQNMINFAPNI---YILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFG 978
Db       1000 GEQNMITMTPSVIATYYLDATGQWENLGVDRRTEAIKQIMTGYAQQMVYKKADHSYAAFT 1059

Qy        979 NYDPSGSTWLSAFVLRCFLEADPYI-DIDQNVLHRTYTWL-KGHQKSNGEFWDPGRVIHS 1036
Db       1060 N--RASSSWLTAYVVKVLAMASNMVKDISHEIICGGVKWLILNRQQPDGVFKENAPVIHG 1117

Qy       1037 ELQGGNKS---PVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRG----------I 1083
Db       1118 EMLGGTKGAEPEASLTAFIVTALLESRSV-----CKEQINILDSSINKATDYLLKKYEKL 1172

Qy       1084 SDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIE 1143
Db       1173 QRPYTTALTAYALA------AADRLNDDRVLMAASTGRNRW--------EEYNARTHNIE 1218

Qy       1144 VAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTER- 1202
Db       1219 GTSYALLALLKMKKFAEVGPVVRWLIDQKYYGGTYGQTQATVMVFQALAEYEIQMPTHQD 1278

Qy       1203 TNIQVTVTGPSSPSP-------LAVVQPMAV--------NISANGFGFAICQLNVVYNVK 1247
Db       1279 LNLDISIKLPEREVPERYSINDRNAVQARTVETKLNEDFTVSASGDGKATMTILTVYNAQ 1338

Qy       1248 ASGSSRRRRSIQNQEAFDLDVAV-------KENKDDLNHVDLNVCTSFSGPGRSGMALME 1300
Db       1339 L----REDANVCNK--FHLDVSVENVELNLKQAKGGKAALRLKICTRYLGEVDSTMTIID 1392

Qy       1301 VNLLSGFMVPSEAI-----SLSETVKKVEYDH-----GKLNLYLDSVNETQ-FCVNIPAV 1349
Db       1393 ISMLTGFFPDAEDLKRLSNGVDRYISKFEIDNNMAQKGTVVIYLDKVSHSEDECLHFKIH 1452

Qy       1350 RNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHH 1409
Db       1453 KHFEVGFIQPGSVKVYSYYNLDEQCTKFYHPDKETGLLNKICHGNICRCAEETCSLLNQQ 1512

Qy       1410 --------------SSVIFIFCFKLL----------YFMEL 1426
Db       1513 KKIDLQLRIQKACAQNVDYVYKTKLLRIEEKDGNDIYFMDV 1553
```

RESULT 3
US-08-793-126-1
; Sequence 1, Application US/08793126
; Patent No. 5849297
; GENERAL INFORMATION:
; APPLICANT: Harrison, Richard Alexander
; APPLICANT: Farries, Charles Timothy
; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: United States of America
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,126
; FILING DATE: 07-FEB-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 102286.377
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 526-6000
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1663 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-793-126-1

Query Match 12.7%; Score 931; DB 2; Length 1663;
Best Local Similarity 23.0%; Pred. No. 3.6e-68;
Matches 376; Conservative 286; Mismatches 615; Indels 356; Gaps 62;

```
Qy          3 GPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIR-------------PGGNVTIGVELL 49
Db          6 GPSLL----LLLLTHLPLALG-SPMYSIITPNILRLESEETMVLEAHDAQGDVPVTVTVH 60
```

```
Qy     50 EH------CPSQVTVKAELLKTASNLTVSV-LEAEGVFEKGSFKTLTLPSLPLNSADEIY 102
Db     61 DFPGKKLVLSSEKTVLTPATNHMGNVTFTIPANREFKSEKGRNKFVTVQA---TFGTQVV 117

Qy    103 ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKT 162
Db    118 EKVVLVSLQSGYL----------------FIQTDKTIYTPGSTVLYRIFTVNHKLLPVGR 161

Qy    163 SLNILIKDPKSNLIQQ-WLSQQSDLGVISKTFQLSSHPILGDWSIQVQVND---QTYYQS 218
Db    162 TVMVNIENPEGIPVKQDSLSSQNQLGVLPLSWDIPELVNMGQWKIRAYYENSPQQVFSTE 221

Qy    219 FQVSEYVLPKFEVTLQ--TPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTF-------- 268
Db    222 FEVKEYVLPSFEVIVEPTEKFYYIYNEKGLEVTITARFLYGKKVEGTAFVIFGIQDGEQR 281

Qy    269 ---------LPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPG 319
Db    282 ISLPESLKRIPIEDGSGEVVLSRKVLLDGVQNPRAEDLVGKSLYVSA------------- 328

Qy    320 PVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGN- 378
Db    329 -----TVILHSGSDMVQAERSGIPIVTSPYQIHFTKTPKYFKPGMPFDLMVFVTNPDGSP 383

Qy    379 --QLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILE 436
Db    384 AYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQK-----------PLSITVRTKKQELS 432

Qy    437 DSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKT------RDENIKVGSPFELVVSGN 490
Db    433 EAEQ--------ATRTMQALPYSTVGNSNNYLHLSVLRTELRPGETLNVNFLLRMDRAHE 484

Qy    491 KRLKELSYMVVSRGQLVAVGK------QNSTMFSLTPENSWTPKACVIVYY----IEDDG 540
Db    485 AKIRYYTYLIMNKGRLLKAGRQVREPGQDLVVLPLSITTDFIPSFRLVAYYTLIGASGQR 544

Qy    541 EIISDVLKIPVQ------LVFKNKIKLYWSKVKAE-----PSEKVSLRISVTQPDSIVGI 589
Db    545 EVVADSVWVDVKDSCVGSLVVKSG--------QSEDRQPVPGQQMTLKIEGDHGARVV-L 595

Qy    590 VAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDA----- 644
Db    596 VAVDKGVFVLNKKNKLTQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSSSGQQTA 655

Qy    645 ---------------------------------NLTKDYIDGVYDN-----AEYAERFME 666
Db    656 QRAELQCPQPAARRRRSVQLTEKRMDKVGKYPKELRKCCEDGMRENPMRFSCQRRTRFIS 715

Qy    667 ENE----------GHIVDIH----------------DFSLGSSPHV--RKHFPETWIW-- 696
Db    716 LGEACKKVFLDCCNYITELRRQHARASHLGLARSNLDEDIIAEENIVSRSEFPESWLWNV 775

Qy    697 ----------LDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQ 746
Db    776 EDLKEPPKNGISTKL-------MNIFLKDSITTWEILAVSMSDKKGI-CVADPFEVTVMQ 827

Qy    747 PFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIE--KSDKFDILMTSSEINATGH 804
Db    828 DFFIDLRLPYSVVRNEQVEIRAVLYNY-RQNQELKVRVELLHNPAFCSLATTKR----RH 882

Qy    805 QQTLLVPSEDGATVLFPIRP--THLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQ 862
Db    883 QQTITIPPKSSLSVPYVIVPLKTGLQEVEVKAAVYHHFISDGVRKSLKVVPEGIRMNKTV 942

Qy    863 SILLDLTDNRL--QSTLKTLSFSFPPNTV------TGSER---VQITAIGDVLGPSING- 910
Db    943 AV-RTLDPERLGREGVQKE---DIPPADLSDQVPDTESETRILLQGTPVAQMTEDAVDAE 998

Qy    911 -LASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEK---ALSFMRQGYQRELL 966
Db    999 RLKHLIVTPSGCGEQNMIGMTPTVIAVHYLDETEQWEKFGLEKRQGALELIKKGYTQQLA 1058

Qy    967 YQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWL-KGHQKSNG 1025
Db   1059 FRQPSSAFAAFVKRAP--STWLTAYVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDG 1116

Qy   1026 EFWDPGRVIHSELQGG----NKSPVTLTAYIVTSLLGYRKYQPNIDV-QESIH------- 1073
Db   1117 VFQEDAPVIHQEMIGGLRNNNEKDMALTAFVLISL------QEAKDICEEQVNSLPGSIT 1170

Qy   1074 ----FLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSES 1129
Db   1171 KAGDFLEANY-MNLQRSYTVAIAGYALAQMGRLKGPLLNKFLT--------------TA 1214

Qy   1130 KLSDSWQ---PRSLDIEVAAYALLSHFLQFQTSEGI-PIMRWLSRQRNSLGGFASTQDTT 1185
Db   1215 KDKNRWEDPGKQLYNVEATSYALLA-LLQLKDFDFVPPVVRWLNEQRYYGGGYGSTQATF 1273

Qy   1186 VALKALSEFAA-LMNTERTNIQVTVTGPSSPSPLA--------------VVQPMAVNIS 1229
Db   1274 MVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITHRIHWESASLLRSEETKENEGFTVT 1333

Qy   1230 ANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAV-------KENKDDLNHVDL 1282
Db   1334 AEGKGQGTLSVVTMYHAKAKDQLTCNK-------FDLKVTIKPAPETEKRPQDAKNTMIL 1386

Qy   1283 NVCTSFSGPGRSGMALMEVNLLSGFMVPSE-----AISLSETVKKVEY-----DHGKLNL 1332
Db   1387 EICTRYRGDQDATMSILDISMMTGFAPDTDDLKQLANGVDRYISKYELDKAFSDRNTLII 1446

Qy   1333 YLDSVNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEV---KLSSCD 1388
Db   1447 YLDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEESCTRFYHPEKEDGKLNK-- 1504

Qy   1389 LCSDVQGCRPCED 1401
Db   1505 LCRD-ELCRCAEE 1516
```

RESULT 4
US-09-132-271-1
; Sequence 1, Application US/09132271
; Patent No. 6221657
; GENERAL INFORMATION:
; APPLICANT: Harrison, Richard Alexander
; APPLICANT: Farries, Charles Timothy
; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: United States of America
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/132,271
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,126
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 102286.377
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 526-6000
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```
;       LENGTH:  1663 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-132-271-1

  Query Match              12.7%;  Score 931;  DB 3;  Length 1663;
  Best Local Similarity    23.0%;  Pred. No. 3.6e-68;
  Matches  376;  Conservative  286;  Mismatches  615;  Indels  356;  Gaps   62;

Qy          3 GPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIR-------------PGGNVTIGVELL 49
Db          6 GPSLL----LLLLTHLPLALG-SPMYSIITPNILRLESEETMVLEAHDAQGDVPVTVTVH 60

Qy         50 EH------CPSQVTVKAELLKTASNLTVSV-LEAEGVFEKGSFKTLTLPSLPLNSADEIY 102
Db         61 DFPGKKLVLSSEKTVLTPATNHMGNVTFTIPANREFKSEKGRNKFVTVQA---TFGTQVV 117

Qy        103 ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKT 162
Db        118 EKVVLVSLQSGYL----------------FIQTDKTIYTPGSTVLYRIFTVNHKLLPVGR 161

Qy        163 SLNILIKDPKSNLIQQ-WLSQQSDLGVISKTFQLSSHPILGDWSIQVQVND---QTYYQS 218
Db        162 TVMVNIENPEGIPVKQDSLSSQNQLGVLPLSWDIPELVNMGQWKIRAYYENSPQQVFSTE 221

Qy        219 FQVSEYVLPKFEVTLQ--TPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTF-------- 268
Db        222 FEVKEYVLPSFEVIVEPTEKFYYIYNEKGLEVTITARFLYGKKVEGTAFVIFGIQDGEQR 281

Qy        269 ---------LPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPG 319
Db        282 ISLPESLKRIPIEDGSGEVVLSRKVLLDGVQNPRAEDLVGKSLYVSA------------- 328

Qy        320 PVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGN- 378
Db        329 -----TVILHSGSDMVQAERSGIPIVTSPYQIHFTKTPKYFKPGMPFDLMVFVTNPDGSP 383

Qy        379 --QLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILE 436
Db        384 AYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQK-----------PLSITVRTKKQELS 432

Qy        437 DSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKT------RDENIKVGSPFELVVSGN 490
Db        433 EAEQ--------ATRTMQALPYSTVGNSNNYLHLSVLRTELRPGETLNVNFLLRMDRAHE 484

Qy        491 KRLKELSYMVVSRGQLVAVGK------QNSTMFSLTPENSWTPKACVIVYY----IEDDG 540
Db        485 AKIRYYTYLIMNKGRLLKAGRQVREPGQDLVVLPLSITTDFIPSFRLVAYYTLIGASGQR 544

Qy        541 EIISDVLKIPVQ------LVFKNKIKLYWSKVKAE-----PSEKVSLRISVTQPDSIVGI 589
Db        545 EVVADSVWVDVKDSCVGSLVVKSG--------QSEDRQPVPGQQMTLKIEGDHGARVV-L 595

Qy        590 VAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDA----- 644
Db        596 VAVDKGVFVLNKKNKLTQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSSSGQQTA 655

Qy        645 ---------------------------------NLTKDYIDGVYDN-----AEYAERFME 666
Db        656 QRAELQCPQPAARRRRSVQLTEKRMDKVGKYPKELRKCCEDGMRENPMRFSCQRRTRFIS 715

Qy        667 ENE----------GHIVDIH----------------DFSLGSSPHV--RKHFPETWIW-- 696
Db        716 LGEACKKVFLDCCNYITELRRQHARASHLGLARSNLDEDIIAEENIVSRSEFPESWLWNV 775

Qy        697 ----------LDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQ 746
Db        776 EDLKEPPKNGISTKL-------MNIFLKDSITTWEILAVSMSDKKGI-CVADPFEVTVMQ 827

Qy        747 PFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIE--KSDKFDILMTSSEINATGH 804
Db        828 DFFIDLRLPYSVVRNEQVEIRAVLYNY-RQNQELKVRVELLHNPAFCSLATTKR----RH 882

Qy        805 QQTLLVPSEDGATVLFPIRP--THLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQ 862
Db        883 QQTITIPPKSSLSVPYVIVPLKTGLQEVEVKAAVYHHFISDGVRKSLKVVPEGIRMNKTV 942

Qy        863 SILLDLTDNRL--QSTLKTLSFSFPPNTV------TGSER---VQITAIGDVLGPSING- 910
Db        943 AV-RTLDPERLGREGVQKE---DIPPADLSDQVPDTESETRILLQGTPVAQMTEDAVDAE 998

Qy        911 -LASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEK---ALSFMRQGYQRELL 966
Db        999 RLKHLIVTPSGCGEQNMIGMTPTVIAVHYLDETEQWEKFGLEKRQGALELIKKGYTQQLA 1058

Qy        967 YQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWL-KGHQKSNG 1025
Db       1059 FRQPSSAFAAFVKRAP--STWLTAYVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDG 1116

Qy       1026 EFWDPGRVIHSELQGG----NKSPVTLTAYIVTSLLGYRKYQPNIDV-QESIH------- 1073
Db       1117 VFQEDAPVIHQEMIGGLRNNNEKDMALTAFVLISL------QEAKDICEEQVNSLPGSIT 1170

Qy       1074 ----FLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSES 1129
Db       1171 KAGDFLEANY-MNLQRSYTVAIAGYALAQMGRLKGPLLNKFLT---------------TA 1214

Qy       1130 KLSDSWQ---PRSLDIEVAAYALLSHFLQFQTSEGI-PIMRWLSRQRNSLGGFASTQDTT 1185
Db       1215 KDKNRWEDPGKQLYNVEATSYALLA-LLQLKDFDFVPPVVRWLNEQRYYGGGYGSTQATF 1273

Qy       1186 VALKALSEFAA-LMNTERTNIQVTVTGPSSPSPLA---------------VVQPMAVNIS 1229
Db       1274 MVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITHRIHWESASLLRSEETKENEGFTVT 1333

Qy       1230 ANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAV-------KENKDDLNHVDL 1282
Db       1334 AEGKGQGTLSVVTMYHAKAKDQLTCNK-------FDLKVTIKPAPETEKRPQDAKNTMIL 1386

Qy       1283 NVCTSFSGPGRSGMALMEVNLLSGFMVPSE-----AISLSETVKKVEY-----DHGKLNL 1332
Db       1387 EICTRYRGDQDATMSILDISMMTGFAPDTDDLKQLANGVDRYISKYELDKAFSDRNTLII 1446

Qy       1333 YLDSVNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEV---KLSSCD 1388
Db       1447 YLDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEESCTRFYHPEKEDGKLNK-- 1504

Qy       1389 LCSDVQGCRPCED 1401
Db       1505 LCRD-ELCRCAEE 1516
```

```
RESULT 5
US-09-142-334-22
; Sequence 22, Application US/09142334
; Patent No. 6268485
; GENERAL INFORMATION:
;  APPLICANT: Farries, Timothy C.
;  APPLICANT: Harrison, Richard A.
;  TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase
;  FILE REFERENCE: 4-30443/A/IMU/PCT
;  CURRENT APPLICATION NUMBER: US/09/142,334
;  CURRENT FILING DATE: 1999-04-15
;  EARLIER APPLICATION NUMBER: PCT/GB97/00603
;  EARLIER FILING DATE: 1997-03-04
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 1663
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-142-334-22

  Query Match              12.7%;  Score 931;  DB 3;  Length 1663;
```

```
Best Local Similarity   23.0%;  Pred. No. 3.6e-68;
Matches  376;  Conservative  286;  Mismatches  615;  Indels  356;  Gaps   62;

Qy          3 GPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIR-------------PGGNVTIGVELL 49
              || ||    || :      ||:    |  : :  |  |:|              |:|  :  |  :
Db          6 GPSLL----LLLLTHLPLALG-SPMYSIITPNILRLESEETMVLEAHDAQGDVPVTVTVH 60

Qy         50 EH------CPSQVTVKAELLKTASNLTVSV-LEAEGVFEKGSFKTLTLPSLPLNSADEIY 102
              :         |: ||          |:|  ::          |||   |  :|:  :          ::
Db         61 DFPGKKLVLSSEKTVLTPATNHMGNVTFTIPANREFKSEKGRNKFVTVQA---TFGTQVV 117

Qy        103 ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKT 162
              |    |     |     |                 ||||||  :|  |   | :||  |:       |
Db        118 EKVVLVSLQSGYL----------------FIQTDKTIYTPGSTVLYRIFTVNHKLLPVGR 161

Qy        163 SLNILIKDPKSNLIQQ-WLSQQSDLGVISKTFQLSSHPILGDWSIQVQVND---QTYYQS 218
              :: :  |::|:    ::|  ||  |: |||:  :: :        :|  |  |:     :    |  :
Db        162 TVMVNIENPEGIPVKQDSLSSQNQLGVLPLSWDIPELVNMGQWKIRAYYENSPQQVFSTE 221

Qy        219 FQVSEYVLPKFEVTLQ--TPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTF-------- 268
              |:| ||||| ||| ::       |       |  |  |   ||||::  |||  |:|     :  |
Db        222 FEVKEYVLPSFEVIVEPTEKFYYIYNEKGLEVTITARFLYGKKVEGTAFVIFGIQDGEQR 281

Qy        269 ---------LPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPG 319
                       :|:    :  :::  :|     |       |       |:  :|  |:
Db        282 ISLPESLKRIPIEDGSGEVVLSRKVLLDGVQNPRAEDLVGKSLYVSA------------- 328

Qy        320 PVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGN- 378
                    |  :    |  :  :  :     :         |   |   |              ||  :  |     |  ||    ||:
Db        329 -----TVILHSGSDMVQAERSGIPIVTSPYQIHFTKTPKYFKPGMPFDLMVFVTNPDGSP 383

Qy        379 --QLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILE 436
                ::  :   :    |  ::||  :         |   :    :||                |  |  |  :  :     |
Db        384 AYRVPVAVQGEDTVQSLTQGDGVAKLSINTHPSQK-----------PLSITVRTKKQELS 432

Qy        437 DSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKT------RDENIKVGSPFELVVSGN 490
              :: :              :::   |:              |     |  |:  |                  |  :  |            :    :
Db        433 EAEQ--------ATRTMQALPYSTVGNSNNYLHLSVLRTELRPGETLNVNFLLRMDRAHE 484

Qy        491 KRLKELSYMVVSRGQLVAVGK------QNSTMFSLTPENSWTPKACVIVYY----IEDDG 540
              :::    :|::::|:|:    |:             |:   :    |:        :   :|    :: ||
Db        485 AKIRYYTYLIMNKGRLLKAGRQVREPGQDLVVLPLSITTDFIPSFRLVAYYTLIGASGQR 544

Qy        541 EIISDVLKIPVQ------LVFKNKIKLYWSKVKAE-----PSEKVSLRISVTQPDSIVGI 589
              |:::|  :  :  |:                ||  |:                 ::|     | ::::|:|                  :|   :
Db        545 EVVADSVWVDVKDSCVGSLVVKSG--------QSEDRQPVPGQQMTLKIEGDHGARVV-L 595

Qy        590 VAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDA----- 644
              |||||  |  ::|     |   :|        :      :|   :   |      |       :     ||  :  ||      :  :
Db        596 VAVDKGVFVLNKKNKLTQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSSSGQQTA 655

Qy        645 --------------------------------NLTKDYIDGVYDN-----AEYAERFME 666
                                              |  |     ||:  :|             :        ||:
Db        656 QRAELQCPQPAARRRRSVQLTEKRMDKVGKYPKELRKCCEDGMRENPMRFSCQRRTRFIS 715

Qy        667 ENE----------GHIVDIH----------------DFSLGSSPHV--RKHFPETWIW-- 696
                |               :|  ::                         |    :  :    ::    |     |||:|:|
Db        716 LGEACKKVFLDCCNYITELRRQHARASHLGLARSNLDEDIIAEENIVSRSEFPESWLWNV 775

Qy        697 ----------LDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQ 746
                          :  |    :       :  |||:|              :|    |:             |   |:     |
Db        776 EDLKEPPKNGISTKL-------MNIFLKDSITTWEILAVSMSDKKGI-CVADPFEVTVMQ 827

Qy        747 PFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIE--KSDKFDILMTSSEINATGH 804
               ||| |  |||||:|  |:     :    ::||   :        |:||  |      |     |  |:                |
Db        828 DFFIDLRLPYSVVRNEQVEIRAVLYNY-RQNQELKVRVELLHNPAFCSLATTKR----RH 882

Qy        805 QQTLLVPSEDGATVLFPIRP--THLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQ 862
              |||:  :|  :           :|  :  |  |     |  |  |:  :                       ||  |  :  :   |     |||    :  :
Db        883 QQTITIPPKSSLSVPYVIVPLKTGLQEVEVKAAVYHHFISDGVRKSLKVVPEGIRMNKTV 942

Qy        863 SILLDLTDNRL--QSTLKTLSFSFPPNTV------TGSER---VQITAIGDVLGPSING- 910
              ::     |      ||    :      |              ||    :             |  ||          :|  |    :    :      :::
Db        943 AV-RTLDPERLGREGVQKE---DIPPADLSDQVPDTESETRILLQGTPVAQMTEDAVDAE 998

Qy        911 -LASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEK---ALSFMRQGYQRELL 966
                |   ||      |  |||||||        |      :     :  ||  :  :|                  ||          ||      :::||   ::|
Db        999 RLKHLIVTPSGCGEQNMIGMTPTVIAVHYLDETEQWEKFGLEKRQGALELIKKGYTQQLA 1058

Qy        967 YQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWL-KGHQKSNG 1025
              :::        :|:||         |         ||||:|:|::   |      |        |  ||     ||                ||        || :|
Db       1059 FRQPSSAFAAFVKRAP--STWLTAYVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDG 1116

Qy       1026 EFWDPGRVIHSELQGG----NKSPVTLTAYIVTSLLGYRKYQPNIDV-QESIH------- 1073
                 |         :     ||  |:  ||           |:     :   |||:::  ||              |        |:  :|  ::
Db       1117 VFQEDAPVIHQEMIGGLRNNNEKDMALTAFVLISL------QEAKDICEEQVNSLPGSIT 1170

Qy       1074 ----FLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSES 1129
                  |||:  :           :    :||:|:        |||:  :|     |                   ||                         :
Db       1171 KAGDFLEANY-MNLQRSYTVAIAGYALAQMGRLKGPLLNKFLT--------------TA 1214

Qy       1130 KLSDSWQ---PRSLDIEVAAYALLSHFLQFQTSEGI-PIMRWLSRQRNSLGGFASTQDTT 1185
                |      :  |:      :      ::|      :||||:      ||  :      :     :    |::|||:  ||               ||:  |||   |
Db       1215 KDKNRWEDPGKQLYNVEATSYALLA-LLQLKDFDFVPPVVRWLNEQRYYGGGYGSTQATF 1273

Qy       1186 VALKALSEFAA-LMNTERTNIQVTVTGPSSPSPLA---------------VVQPMAVNIS 1229
                :    :||:::      :  :      |:  |::      ||    |   :                                            :        ::
Db       1274 MVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITHRIHWESASLLRSEETKENEGFTVT 1333

Qy       1230 ANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAV-------KENKDDLNHVDL 1282
                ||    |       |         :       :|:  ||               |         |||  |  :              |     :|      |      :   |
Db       1334 AEGKGQGTLSVVTMYHAKAKDQLTCNK-------FDLKVTIKPAPETEKRPQDAKNTMIL 1386

Qy       1283 NVCTSFSGPGRSGMALMEVNLLSGFMVPSE-----AISLSETVKKVEY-----DHGKLNL 1332
                :||  :         :    |:::::::::||   ::                      :        :     :     |    |     :
Db       1387 EICTRYRGDQDATMSILDISMMTGFAPDTDDLKQLANGVDRYISKYELDKAFSDRNTLII 1446

Qy       1333 YLDSVNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEV---KLSSCD 1388
              |||  |:  ::   |:             :   |   |      |       :|   :         ||                    |     |:  |          ||:
Db       1447 YLDKVSHSEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEESCTRFYHPEKEDGKLNK-- 1504

Qy       1389 LCSDVQGCRPCED 1401
              ||  |   :   ||    |:
Db       1505 LCRD-ELCRCAEE 1516
```

RESULT 6
US-08-487-283A-2
; Sequence 2, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott
; APPLICANT: Nye, Steven H.
; APPLICANT: Rollins, Scott
; APPLICANT: Rother, Russell P.
; APPLICANT: Springhorn, Jeremy P.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: Thomas, Thomas C.
; APPLICANT: Wilkins, James A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
; TITLE OF INVENTION: OF INFLAMMATORY DISEASES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seth A. Fidel
; STREET: 25 Science Park (Alexion)
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06511
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.4Mb storage
; COMPUTER: Macintosh Cetris 610
; OPERATING SYSTEM: System 7
; SOFTWARE: WordPerfect 3.0

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,283A
;      FILING DATE:  June 7, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/236,208
;      FILING DATE:  02-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seth A. Fidel.
;      REGISTRATION NUMBER:  38,449
;      REFERENCE/DOCKET NUMBER:  ALX-152.1 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (203)776-1790
;      TELEFAX:  (203)772-3655
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1676 Amino Acids
;      TYPE:  Amino Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:
;      DESCRIPTION:  Pro-C5 Polytpeptide
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    PUBLICATION INFORMATION:
;      AUTHORS:  Haviland, D.L.
;      AUTHORS:  Haviland, J.C.
;      AUTHORS:  Fleischer, D.T.
;      AUTHORS:  Hunt, A.
;      AUTHORS:  Wetsel, R.A.
;      TITLE:  Complete cDNA Sequence of Human
; Patent No. 6355245
;      TITLE:  Complement Pro-C5
;      JOURNAL:  Journal of Immunology
;      VOLUME:  146
;      PAGES:  362-368
;      DATE:  1991
US-08-487-283A-2

  Query Match             12.5%;  Score 921;  DB 4;  Length 1676;
  Best Local Similarity   24.0%;  Pred. No. 2.5e-67;
  Matches  379;  Conservative  297;  Mismatches  621;  Indels  284;  Gaps   67;

Qy         28 FLVTAPGIIRPGGNVTIGVEL---LEHCPSQVTVKAELLKTASNLTVSV-LEAEGVFEKG 83
Db         23 YVISAPKIFRVGASENIVIQVYGYTEAFDATISIKSYPDKKFSYSSGHVHLSSENKFQNS 82
Qy         84 SFKTLTLPSLP--LNSADEIYELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYK 141
Db         83 AILTIQPKQLPGGQNPVSYVY-LEVVSKH-----FSKSKRMPITYDNGFLFIHTDKPVYT 136
Qy        142 PKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVIS-KTFQLSSHPI 200
Db        137 PDQSVKVRVYSLNDDLKPAKRETVLTFIDPEGSEVDM-VEEIDHIGIISFPDFKIPSNPR 195
Qy        201 LGDWSIQVQVNDQ-----TYYQSFQVSEYVLPKFEVTLQTPLYCSM---NSKHLNGTITA 252
Db        196 YGMWTIKAKYKEDFSTTGTAY--FEVKEYVLPHFSVSIE-PEYNFIGYKNFKNFEITIKA 252
Qy        253 KYTYGKPV-KGDVTLTFLPLSFWGKKKNI-----------TKTFKINGSANFSFNDEEM 299
Db        253 RYFYNKVVTEADVYITF------GIREDLKDDQKEMMQTAMQNTMLINGIAQVTFDSE-- 304
Qy        300 KNVMDSSNGLSEYL--DLSSPGPVEILTTVTESVTGISRNVS-TNVFFKQHDYIIEFFDY 356
Db        305 ----TAVKELSYYSLEDLNNK-YLYIAVTVIESTGGFSEEAEIPGIKYVLSPYKLNLVAT 359
Qy        357 TTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI-----TVTQRNYTEYWSGSNSGNQK 411
Db        360 PLFLKPGIPYPIKVQV------KDSLDQLVGGVPVILNAQTIDVNQETSDLDPSKSVTRV 413
Qy        412 MEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLK 471
Db        414 DDGVASFVLNLP-SGVTVLEFNVKTDAPDLPEENQAREGYRAIAYSSL--SQSYLYIDWT 470
Qy        472 TRDENIKVGSPFELVVSGNK----RLKELSYMVVSRGQLVAVGKQNS------TMFSLTP 521
Db        471 DNHKALLVGEHLNIIVTPKSPYIDKITHYNYLILSKGKIIHFGTREKFSDASYQSINIPV 530
Qy        522 ENSWTPKACVIVYYI---EDDGEIISDVLKIPVQLVFKNKIKLYWS--KVKAEPSEKVSL 576
Db        531 TQNMVPSSRLLVYYIVTGEQTAELVSDSVWLNIEEKCGNQLQVHLSPDADAYSPGQTVSL 590
Qy        577 RISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQEC 636
Db        591 NMA-TGMDSWVALAAVDSAVYGVQRGAKKPLERVFQFLEKSDLGCGAGGGLNNANVFHLA 649
Qy        637 GLWVLTDAN--------------------------LTKDYIDGV-----YDNA------ 658
Db        650 GLTFLTNANADDSQENDEPCKEILRPRRTLQKKIEEIAAKYKHSVVKKCCYDGACVNNDE 709
Qy        659 ---EYAERF------------------MEENEGHIVDIHDFSLG----------SSPHV 686
Db        710 TCEQRAARISLGPRCIKAFTECCVVASQLRANISH----KDMQLGRLHMKTLLPVSKPEI 765
Qy        687 RKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQ 746
Db        766 RSYFPESWLW-EVHLVPR-RKQLQFALPDSLTTWEIQGIGIS-NTGICVADT-VKAKVFK 821
Qy        747 PFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQ 806
Db        822 DVFLEMNIPYSVVRGEQIQLKGTVYNYRTSGMQFCVKMSAVEG----ICTSESPVIDHQG 877
Qy        807 T-------LLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKS 859
Db        878 TKSSKCVRQKVEGSSSHLVTFTVLPLEIGLHNINFSLETWFGKEILVKTLRVVPEGVKRE 937
Qy        860 YSQSILLDLTDNRLQSTL---KTLSFSFPPNTVTGSERVQITA-----IGDVLGP--SIN 909
Db        938 SYSGVTLD--PRGIYGTISRRKEFPYRIPLDLVPKTEIKRILSVKGLLVGEILSAVLSQE 995
Qy        910 GLASLIRMPYGCGEQNMINFAPNIYILDY--------------LTKKKQLTDNLKEKALS 955
Db        996 GINILTHLPKGSAEAELMSVVPVFYVFHYLETGNHWNIFHSDPLIEKQKLKKKLKEGMLS 1055
Qy        956 FMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYT 1015
Db       1056 IMS--------YRNADYSYSVWKG--GSASTWLTAFALRVLGQVNKYVEQNQNSICNSLL 1105
Qy       1016 WL-KGHQKSNGEFWDPGRVIHSELQG-----GNKSPVTLTAYIVTSLLGYRK---YQPNI 1066
Db       1106 WLVENYQLDNGSFKENSQYQPIKLQGTLPVEARENSLYLTAFTV---IGIRKAFDICPLV 1162
Qy       1067 DVQESI----HFLESEFSRGISDNYTLALITYALSSVGS---PKAKEALNMLTWRAEQEG 1119
Db       1163 KIDTALIKADNFL-LENTLPAQSTFTLAISAYAL-SLGDKTHPQFRSIVSALKREALVKG 1220
Qy       1120 G---MQFWVSS----ESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQR 1172
Db       1221 NPPIYRFWKDNLQHKDSSVPNTGTARM--VETTAYALLTSLNLKDINYVNPVIKWLSEEQ 1278
Qy       1173 NSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPS--------------PL 1218
Db       1279 RYGGGFYSTQDTINAIEGLTEYSLLVKQLRLSMDIDVSYKHKGALHNYKMTDKNFLGRPV 1338
Qy       1219 AVVQPMAVN---ISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQ-EAFDLDVAVKENK 1274
Db       1339 EVL----LNDDLIVSTGFGSGLATVHVTTVVHKTSTSEEVCSFYLKIDTQDIEASHYRGY 1394
Qy       1275 DDLNHVDLNVCTSFSGPGR------SGMALMEVNLLSGFMVPSEAI-SLSETVKKVEYDH 1327
Db       1395 GNSDYKRIVACASYK-PSREESSSGSSHAVMDISLPTGISANEEDLKALVEGVDQLFTDY 1453
Qy       1328 ----GKLNLYLDSVNETQF-CVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEV 1382
```

```
Db       1454 QIKDGHVILQLNSIPSSDFLCVRFRIFELFEVGFLSPATFTVYEYHRPDKQCTMFYST-- 1511

Qy       1383 KLSSCDLCSDVQGCRPCEDGA 1403
                      |:::  :  ||  |
Db       1512 --------SNIKIQKVCEGAA 1524

RESULT 7
US-08-662-227-2
; Sequence 2, Application US/08662227
; Patent No. 5922320
;  GENERAL INFORMATION:
;    APPLICANT:  VOGEL, CARL-WILHELM
;    APPLICANT:  BREDEHORST, REINHORST
;    APPLICANT:  KOCK, MICHAEL
;    APPLICANT:  FRITZINGER, DAVID
;    TITLE OF INVENTION:  RECOMBINANT PROCVF
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/662,227
;      FILING DATE:  14-JUN-1996
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-0107-0X
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1642 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-662-227-2

  Query Match             12.2%;  Score 894;  DB 2;  Length 1642;
  Best Local Similarity   22.5%;  Pred. No. 4.4e-65;
  Matches  361;  Conservative  295;  Mismatches  672;  Indels  276;  Gaps   56;

Qy         18 AALAV-----APGPRFLVTAPGIIRPGGNVTIGVELL-EHCPSQVTVKA----ELLKTAS 67
              |||  :      :  |  :  :   |  ::|          | ||    :  |  |: :          ||
Db         10 AALLIGFPGSSHGALYTLITPAVLRTDTEEQILVEAHGDSTPKQLDIFVHDFPRKQKTLF 69

Qy         68 NLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETK 127
                  | :  |  |:         |: :|:  :::          |    |   ::          ||:::
Db         70 QTRVDMNPAGGMLVT---PTIEIPAKEVSTDSRQNQYVVVQVTGPQVRLEKVVLLSYQSS 126

Qy        128 RISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLG 187
                :||||| :|  |    |  :|: ::   :      :: :  : |:  |:      |    ||
Db        127 --FLFIQTDKGIYTPGSPVLYRVFSMDHNTSKMNKTVIVEFQTPEGILVS---SNSVDLN 181

Qy        188 VISKTFQLSSHPILGDWSI--QVQVNDQTYYQSFQVSEYVLPKFEVTLQTP---LYCSMN 242
                  : |        || |  |   : : : :  |     |  | :|||| ||| ||          |    |
Db        182 FF-WPYNLPDLVSLGTWRIVAKYEHSPENYTAYFDVRKYVLPSFEVRLQPSEKFFYIDGN 240

Qy        243 SKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNI----TKTFKINGSANFSFNDEE 298
                 :: : :|||:| ||: |:|  |        :          ||:|      |:    |:|      :     :

Db        241 -ENFHVSITARYLYGEEVEG-VAFVLFGVKIDDAKKSIPDSLTRIPIIDGDGKATLKRDT 298

Qy        299 MKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNFFKQHDYIIEFFDYTT 358
               ::   : | |   :   :|          :| :|||  : :         : :            | | |
Db        299 FRSRFPNLNELVGHTLYAS------VTVMTESGSDMVVTEQSGIHIVASPYQIHFTKTPK 352

Qy        359 VLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKI 418
                || : :   || ||   ||:                :  |       :             :  |       :    |:
Db        353 YFKPGMPYELTVYVTNPDGSPA--------AHVPVVSEAF------HSMGTTLSDGTAKL 398

Qy        419 NYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIK 478
               :| :              ||   ::       |          :||  |:      :      |:  :          ||
Db        399 ILNIPLNAQ---SLPITVRTNHGDLPRERQATKSMTAIAYQTQGGSGNYLHVAITSTEIK 455

Qy        479 VGS--PFELVVSGN----KRLKELSYMVVSRGQLVAVGK------QNSTMFSLTPENSWT 526
                |      |        |  ||         |::|   :|:::::|::   :||:           ||        :|
Db        456 PGDNLPVNFNVKGNANSLKQIKYFTYLILNKGKIFKVGRQPRRDGQNLVTMNLHITPDLI 515

Qy        527 PKACVIVYYIEDDGEIISDVLKIPVQ------LVFK--NKIKLYWSKVKAEPSEKVSLRI 578
               |         :  ||      :  ||::|  :  :  |:          ||  |     |  |::               |       :  :::
Db        516 PSFRFVAYYQVGNNEIVADSVWVDVKDTCMGTLVVKGDNLIQM--------PGAAMKIKL 567

Qy        579 SVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGL 638
                    |  :  ||:|||||:|  ::|        |:     :     :|     :   |          |          |:     ||:: ||
Db        568 E-GDPGARVGLVAVDKAVYVLNDKYKISQAKIWDTIEKSDFGCTAGSGQNNLGVFEDAGL 626

Qy        639 WVLTDANL-TKD------------------------------------------------ 649
               :  |    || ||
Db        627 ALTTSTNLNTKQRSAAKCPQPANRRRRSSVLLLDSNASKAAEFQDQDLRKCCEDVMHENP 686

Qy        650 --------------------------YIDGVYD-NAEYAERFM--EENE-GHIVDIHDF 678
                                          || || |   |     :|  |:    ::||  |   |   |
Db        687 MGYTCEKRAKYIQEGDACKAAFLECCRYIKGVRDENQRESELFLARDDNEDGFIADSDII 746

Qy        679 SLGSSPHVRKHFPETWIWL------DTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLG 732
               |          |     ||::|:||              :  |             :            :    ||||:||                :      |
Db        747 S-------RSDFPKSWLWLTKDLTEEPNSQGISSKTMSFYLRDSITTWVVLAVSFTPTKG 799

Qy        733 LGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDI 792
               :           |  |::    :   |||   |    :||||::  |:      :        :    ||: :          |:|      :      :       |
Db        800 I-CVAEPYEIRVMKVFFIDLQMPYSVVKNEQVEIRAILHNYVNEDIYVRVELLYNPAF-- 856

Qy        793 LMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLG--EIPITVTALSPTASDAVTQMIL 850
                 |:             ::|         |     |     |     |   :: :         :                 || |    :    :
Db        857 --CSASTKGQRYRQQFPIKALSSRAVPFVIVPLEQGLHDVEIKASVQEALWSDGVRKKLK 914

Qy        851 VKAEGIEKSYSQSILLDLTDNRLQST------LKTLSFSFPPNTVTGSERVQITAIGDVL 904
               |     ||::||            :  ||                    |                        :      |              :        :|        :       ::
Db        915 VVPEGVQKSIVTIVKLDPRAKGVGGTQLEVIKARKLDDRVPDTEIETKIIIQGDPVAQII 974

Qy        905 GPSING--LASLIRMPYGCGEQNMINFAPNI---YILDYLTKKKQLTDNLKEKALSFMRQ 959
               ||:|     |      ||     |   |||||||    |       :          |   ||       : :    |        |   :  :|::  :
Db        975 ENSIDGSKLNHLIITPSGCGEQNMIRMAAPVIATYYLDTTEQWETLGINRRTEAVNQIVT 1034

Qy        960 GYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF-LEADPYIDIDQNVLHRTYTWL- 1017
               ||  ::::|::  |   |::||    |         :  |:||:|:|::    |   :  |                 |            ::               ||
Db       1035 GYAQQMVYKKADHSYAAFTN--RASSSWLTAYVVKVFAMAAKMVAGISHEIICGGVRWLI 1092

Qy       1018 KGHQKSNGEFWDPGRVIHSELQG---GNKSPVTLTAYIVTSLLGYR----KYQPNID--V 1068
                   |:  :|  |  :        |:        :||        |    :        |  |||:|  :  :||       :                |        ::|        :
Db       1093 LNRQQPDGAFKENAPVLSGTMQGGIQGAEEEVYLTAFILVALLESKTICNDYVNSLDSSI 1152

Qy       1069 QESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSE 1128
               ::: :|  :   :: :  :  :             || ||   |||:            :|      ||                                  |                  |
Db       1153 KKATNYLLKKYEK-LQRPYTTALTAYALA------AADQLNDDRVLMAASTGRDHW---- 1201

Qy       1129 SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVAL 1188
                  :  :    :  :||    :|||:              :    :           ||:|||:   |                  :         ||    |  :  |
Db       1202 ----EEYNAHTHNIEGTSYALLALLKMKKFDQTGPIVRWLTDQNFYGETYGQTQATVMAF 1257

Qy       1189 KALSEFAALMNTER-TNIQVTVTGPSSPSPL-----------AVVQP---MAVNISANG 1232
               :||:|:       |    |   :     |:  :|:    |                |:                                 |:                      : ::|:|
Db       1258 QALAEYEIQMPTHKDLNLDITIELPDREVPIRYRINYENALLARTVETKLNQDITVTASG 1317
```

```
Qy     1233 FGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVD------LNVCT 1286
Db     1318 DGKATMTILTFYNAQL----QEKANVCNK--FHLNVSVE--NIHLNAMGAKGALMLKICT 1369

Qy     1287 SFSGPGRSGMALMEVNLLSGFMVPSEAIS-----LSETVKKVEYDHGKLN-----LYLDS 1336
Db     1370 RYLGEVDSTMTIIDISMLTGFLPDAEDLTRLSKGVDRYISRYEVDNNMAQKVAVIIYLNK 1429

Qy     1337 VNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQG 1395
Db     1430 VSHSEDECLHFKILKHFEVGFIQPGSVKVYSYYNLDEKCTKFYHPDKGTGLLNKICIGNV 1489

Qy     1396 CRPCEDGASGSHH--------------HSSVIFIFCFKLLYFME 1425
Db     1490 CRCAGETCSSLNHQERIDVPLQIEKACETNVDYVYKTKLLRIEE 1533
```

RESULT 8
US-09-017-947-2
; Sequence 2, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/662,227
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1642 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-017-947-2

```
  Query Match             12.2%;  Score 894;  DB 4;  Length 1642;
  Best Local Similarity   22.5%;  Pred. No. 4.4e-65;
  Matches 361;  Conservative  295;  Mismatches  672;  Indels  276;  Gaps   56;

Qy       18 AALAV-----APGPRFLVTAPGIIRPGGNVTIGVELL-EHCPSQVTVKA----ELLKTAS 67
Db       10 AALLIGFPGSSHGALYTLITPAVLRTDTEEQILVEAHGDSTPKQLDIFVHDFPRKQKTLF 69

Qy       68 NLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETK 127
Db       70 QTRVDMNPAGGMLVT---PTIEIPAKEVSTDSRQNQYVVVQVTGPQVRLEKVVLLSYQSS 126

Qy      128 RISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLG 187
Db      127 --FLFIQTDKGIYTPGSPVLYRVFSMDHNTSKMNKTVIVEFQTPEGILVS---SNSVDLN 181

Qy      188 VISKTFQLSSHPILGDWSI--QVQVNDQTYYQSFQVSEYVLPKFEVTLQTP---LYCSMN 242
Db      182 FF-WPYNLPDLVSLGTWRIVAKYEHSPENYTAYFDVRKYVLPSFEVRLQPSEKFFYIDGN 240

Qy      243 SKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNI----TKTFKINGSANFSFNDEE 298
Db      241 -ENFHVSITARYLYGEEVEG-VAFVLFGVKIDDAKKSIPDSLTRIPIIDGDGKATLKRDT 298

Qy      299 MKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTT 358
Db      299 FRSRFPNLNELVGHTLYAS------VTVMTESGSDMVVTEQSGIHIVASPYQIHFTKTPK 352

Qy      359 VLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKI 418
Db      353 YFKPGMPYELTVYVTNPDGSPA--------AHVPVVSEAF------HSMGTTLSDGTAKL 398

Qy      419 NYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIK 478
Db      399 ILNIPLNAQ---SLPITVRTNHGDLPRERQATKSMTAIAYQTQGGSGNYLHVAITSTEIK 455

Qy      479 VGS--PFELVVSGN----KRLKELSYMVVSRGQLVAVGK------QNSTMFSLTPENSWT 526
Db      456 PGDNLPVNFNVKGNANSLKQIKYFTYLILNKGKIFKVGRQPRRDGQNLVTMNLHITPDLI 515

Qy      527 PKACVIVYYIEDDGEIISDVLKIPVQ------LVFK--NKIKLYWSKVKAEPSEKVSLRI 578
Db      516 PSFRFVAYYQVGNNEIVADSVWVDVKDTCMGTLVVKGDNLIQM--------PGAAMKIKL 567

Qy      579 SVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGL 638
Db      568 E-GDPGARVGLVAVDKAVYVLNDKYKISQAKIWDTIEKSDFGCTAGSGQNNLGVFEDAGL 626

Qy      639 WVLTDANL-TKD------------------------------------------------ 649
Db      627 ALTTSTNLNTKQRSAAKCPQPANRRRRSSVLLLDSNASKAAEFQDQDLRKCCEDVMHENP 686

Qy      650 ---------------------------YIDGVYD-NAEYAERFM--EENE-GHIVDIHDF 678
Db      687 MGYTCEKRAKYIQEGDACKAAFLECCRYIKGVRDENQRESELFLARDDNEDGFIADSDII 746

Qy      679 SLGSSPHVRKHFPETWIWL------DTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLG 732
Db      747 S-------RSDFPKSWLWLTKDLTEEPNSQGISSKTMSFYLRDSITTWVVLAVSFTPTKG 799

Qy      733 LGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDI 792
Db      800 I-CVAEPYEIRVMKVFFIDLQMPYSVVKNEQVEIRAILHNYVNEDIYVRVELLYNPAF-- 856

Qy      793 LMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLG--EIPITVTALSPTASDAVTQMIL 850
Db      857 --CSASTKGQRYRQQFPIKALSSRAVPFVIVPLEQGLHDVEIKASVQEALWSDGVRKKLK 914

Qy      851 VKAEGIEKSYSQSILLDLTDNRLQST------LKTLSFSFPPNTVTGSERVQITAIGDVL 904
Db      915 VVPEGVQKSIVTIVKLDPRAKGVGGTQLEVIKARKLDDRVPDTEIETKIIIQGDPVAQII 974

Qy      905 GPSING--LASLIRMPYGCGEQNMINFAPNI---YILDYLTKKKQLTDNLKEKALSFMRQ 959
Db      975 ENSIDGSKLNHLIITPSGCGEQNMIRMAAPVIATYYLDTTEQWETLGINRRTEAVNQIVT 1034

Qy      960 GYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF-LEADPYIDIDQNVLHRTYTWL- 1017
Db     1035 GYAQQMVYKKADHSYAAFTN--RASSSWLTAYVVKVFAMAAKMVAGISHEIICGGVRWLI 1092

Qy     1018 KGHQKSNGEFWDPGRVIHSELQG---GNKSPVTLTAYIVTSLLGYR----KYQPNID--V 1068
```

```
Db       1093 LNRQQPDGAFKENAPVLSGTMQGGIQGAEEEVYLTAFILVALLESKTICNDYVNSLDSSI 1152

Qy       1069 QESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSE 1128

Db       1153 KKATNYLLKKYEK-LQRPYTTALTAYALA------AADQLNDDRVLMAASTGRDHW---- 1201

Qy       1129 SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVAL 1188

Db       1202 ----EEYNAHTHNIEGTSYALLALLKMKKFDQTGPIVRWLTDQNFYGETYGQTQATVMAF 1257

Qy       1189 KALSEFAALMNTER-TNIQVTVTGPSSPSPL-----------AVVQP---MAVNISANG 1232

Db       1258 QALAEYEIQMPTHKDLNLDITIELPDREVPIRYRINYENALLARTVETKLNQDITVTASG 1317

Qy       1233 FGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVD------LNVCT 1286

Db       1318 DGKATMTILTFYNAQL----QEKANVCNK--FHLNVSVE--NIHLNAMGAKGALMLKICT 1369

Qy       1287 SFSGPGRSGMALMEVNLLSGFMVPSEAIS-----LSETVKKVEYDHGKLN-----LYLDS 1336

Db       1370 RYLGEVDSTMTIIDISMLTGFLPDAEDLTRLSKGVDRYISRYEVDNNMAQKVAVIIYLNK 1429

Qy       1337 VNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQG 1395

Db       1430 VSHSEDECLHFKILKHFEVGFIQPGSVKVYSYYNLDEKCTKFYHPDKGTGLLNKICIGNV 1489

Qy       1396 CRPCEDGASGSHH-------------HSSVIFIFCFKLLYFME 1425

Db       1490 CRCAGETCSSLNHQERIDVPLQIEKACETNVDYVYKTKLLRIEE 1533
```

RESULT 9
US-08-447-411-45
; Sequence 45, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1642 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-447-411-45

Query Match 12.0%; Score 885; DB 1; Length 1642;
Best Local Similarity 22.4%; Pred. No. 2.5e-64;
Matches 359; Conservative 296; Mismatches 673; Indels 276; Gaps 56;

```
Qy         18 AALAV-----APGPRFLVTAPGIIRPGGNVTIGVELL-EHCPSQVTVKA----ELLKTAS 67

Db         10 AALLIGFPGSSHGALYTLITPAVLRTDTEEQILVEAHGDSTPKQLDIFVHDFPRKQKTLF 69

Qy         68 NLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETK 127

Db         70 QTRVDMNPAGGMLVT---PTIEIPAKEVSTDSRQNQYVVVQVTGPQVRLEKVVLLSYQSS 126

Qy        128 RISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLG 187

Db        127 --FLFIQTDKGIYTPGSPVLYRVFSMDHNTSKMNKTVIVEFQTPEGILVS---SNSVDLN 181

Qy        188 VISKTFQLSSHPILGDWSI--QVQVNDQTYYQSFQVSEYVLPKFEVTLQTP---LYCSMN 242

Db        182 FF-WPYNLPDLVSLGTWRIVAKYEHSPENYTAYFDVRKYVLPSFEVRLQPSEKFFYIDGN 240

Qy        243 SKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNI----TKTFKINGSANFSFNDEE 298

Db        241 -ENFHVSITARYLYGEEVEG-VAFVLFGVKIDDAKKSIPDSLTRIPIIDGDGKATLKRDT 298

Qy        299 MKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTT 358

Db        299 FRSRFPNLNELVGHTLYAS------VTVMTESGSDMVVTEQSGIHIVASPYQIHFTKTPK 352

Qy        359 VLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKI 418

Db        353 YFKPGMPYELTVYVTNPDGSPA--------AHVPVVSEAF------HSMGTTLSDGTAKL 398

Qy        419 NYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIK 478

Db        399 ILNIPLNAQ---SLPITVRTNHGDLPRERQATKSMTAIAYQTQGGSGNYLHVAITSTEIK 455

Qy        479 VGS--PFELVVSGN----KRLKELSYMVVSRGQLVAVGK------QNSTMFSLTPENSWT 526

Db        456 PGDNLPVKFQCEGQCNSLKQIKYFTYLILNKGKIFKVGRQPRRDGQNLVTMNLHITPDLI 515

Qy        527 PKACVIVYYIEDDGEIISDVLKIPVQ------LVFK--NKIKLYWSKVKAEPSEKVSLRI 578

Db        516 PSFRFVAYYQVGNNEIVADSVWVDVKDTCMGTLVVKGDNLIQM--------PGAAMKIKL 567

Qy        579 SVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGL 638

Db        568 E-GDPGARVGLVAVDKAVYVLNDKYKISQAKIWDTIEKSDFGCTAGSGQNNLGVFEDAGL 626

Qy        639 WVLTDANL-TKD------------------------------------------------ 649

Db        627 ALTTSTNLNTKQRSAAKCPQPANRRRRSSVLLLDSNASKAAEFQDQDLRKCCEDVMHENP 686

Qy        650 ---------------------------YIDGVYD-NAEYAERFM--EENE-GHIVDIHDF 678

Db        687 MGYTCEKRAKYIQEGDACKAAFLECCRYIKGVRDENQRESELFLARDDNEDGFIADSDII 746

Qy        679 SLGSSPHVRKHFPETWIWL------DTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLG 732

Db        747 S-------RSDFPKSWLWLTKDLTEEPNSQGISSKTMSFYLRDSITTWVVLAVSFTPTKG 799

Qy        733 LGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDI 792

Db        800 I-CVAEPYEIRVMKVFFIDLQMPYSVVKNEQVEIRAILHNYVNEDIYVRVELLYNPAF-- 856

Qy        793 LMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLG--EIPITVTALSPTASDAVTQMIL 850

Db        857 --CSASTKGQRYRQQFPIKALSSRAVPFVIVPLEQGLHDVEIKASVQEALWSDGVRKKLK 914
```

```
Qy        851 VKAEGIEKSYSQSILLDLTDNRLQST------LKTLSFSFPPNTVTGSERVQITAIGDVL 904
Db        915 VVPEGVQKSIVTIVKLDPRAKGVGGTQLEVIKARKLDDRVPDTEIETKIIIQGDPVAQII 974

Qy        905 GPSING--LASLIRMPYGCGEQNMINFAPNI---YILDYLTKKKQLTDNLKEKALSFMRQ 959
Db        975 ENSIDGSKLNHLIITPSGCGEQNMIRMAAPVIATYYLDTTEQWETLGINRRTEAVNQIVT 1034

Qy        960 GYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF-LEADPYIDIDQNVLHRTYTWL- 1017
Db       1035 GYAQQMVYKKADHSYAAFTN--RASSSWLTAYVVKVFAMAAKMVAGISHEIICGGVRWLI 1092

Qy       1018 KGHQKSNGEFWDPGRVIHSELQG---GNKSPVTLTAYIVTSLLGYR----KYQPNID--V 1068
Db       1093 LNRQQPDGAFKENAPVLSGTMQGGIQGAEEEVYLTAFILVALLESKTICNDYVNSLDSSI 1152

Qy       1069 QESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSE 1128
Db       1153 KKATNYLLKKYEK-LQRPYTTALTAYALA------AADQLNDDRVLMAASTGRDHW---- 1201

Qy       1129 SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVAL 1188
Db       1202 ---EEYNAHTHNIEGTSYALLALLKMKKFDQTGPIVRWLTDQNFYGETYGQTQATVMAF 1257

Qy       1189 KALSEFAALMNTER-TNIQVTVTGPSSPSPL-----------AVVQP---MAVNISANG 1232
Db       1258 QALAEYEIQMPTHKDLNLDITIELPDREVPIRYRINYENALLARTVETKLNQDITVTASG 1317

Qy       1233 FGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVD------LNVCT 1286
Db       1318 DGKATMTILTFYNAQL----QEKANVCNK--FHLNVSVE--NIHLNAMGAKGALMLKICT 1369

Qy       1287 SFSGPGRSGMALMEVNLLSGFMVPSEAIS-----LSETVKKVEYDHGKLN-----LYLDS 1336
Db       1370 RYLGEVDSTMTIIDISMLTGFLPDAEDLTRLSKGVDRYISRYEVDNNMAQKVAVIIYLNK 1429

Qy       1337 VNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQG 1395
Db       1430 VSHSEDECLHFKILKHFEVGFIQPGSVKVYSYYNLDEKCTKFYHPDKGTGLLNKICIGNV 1489

Qy       1396 CRPCEDGASGSHH-------------HSSVIFIFCFKLLYFME 1425
Db       1490 CRCAGETCSSLNHQERIDVPLQIEKACETNVDYVYKTKLLRIEE 1533
```

RESULT 10
US-08-662-227-35
; Sequence 35, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,227
; FILING DATE: 14-JUN-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1648 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-662-227-35

Query Match 12.0%; Score 880; DB 2; Length 1648;
Best Local Similarity 22.7%; Pred. No. 6.6e-64;
Matches 365; Conservative 295; Mismatches 664; Indels 284; Gaps 59;

```
Qy         18 AALAV-----APGPRFLVTAPGIIRPGGNVTIGVELL-EHCPSQVTVKA----ELLKTAS 67
Db         10 AALLIGFPGSSHGALYTLITPAVLRTDTEEQILVEAHGDSTPKQLDIFVHDFPRKQKTLF 69

Qy         68 NLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETK 127
Db         70 QTRVDMNPAGGMLVT---PTIEIPAKEVSTDSRQNQYVVVQVTGPQVRLEKVVLLSYQSS 126

Qy        128 RISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILI--KDPKSNLIQQWLSQQSD 185
Db        127 --FLFIQTDKGIYTPGSPVLYRVFSM--DHHTSKMNKTVIVEFQTPEGILVS---SNSVD 179

Qy        186 LGVISKTFQLSSHPILGDWSI--QVQVNDQTYYQSFQVSEYVLPKFEVTLQTP---LYCS 240
Db        180 LNFF-WPYNLPDLVSLGTWRIVAKYEHSPENYTAYFDVRKYVLPSFEVRLQPSEKFFYID 238

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNI----TKTFKINGSANFSFND 296
Db        239 GN-ENFHVSITARYLYGEEVEG-VAFVLFGVKIDDAKKSIPDSLTRIPIIDGDGKATLKR 296

Qy        297 EEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDY 356
Db        297 DTFRSRFPNLNELVGHTLYAS------VTVMTESGSDMVVTEQSGIHIVASPYQIHFTKT 350

Qy        357 TTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQ 416
Db        351 PKYFKPGMPYELTVYVTNPDGSPA--------AHVPVVSEAF------HSMGTTLSDGTA 396

Qy        417 KINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDEN 476
Db        397 KLILNIPLNAQ---SLPITVRTNHGDLPRERQATKSMTAIAYQTQGGSGNYLHVAITSTE 453

Qy        477 IKVGS--PFELVVSGN----KRLKELSYMVVSRGQLVAVGKQNS-------TM-FSLTPE 522
Db        454 IKPGDNLPVNFNVKGNANSLKQIKYFTYLILNKGKIFKVGRQPRRDGQNLVTMNLHITPD 513

Qy        523 NSWTPKACVIVYYIEDDGEIISDVLKIPVQ------LVFK--NKIKLYWSKVKAEPSEKV 574
Db        514 LIPSPR--FVAYYQVGNNEIVADSVWVDVKDTCMGTLVVKGDNLIQMPGAAMKIKLEGDF 571

Qy        575 SLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQ 634
Db        572 GAR---------VGLVAVDKAVYVLNDKYKISQAKIWDTIEKSDFGCTAGSGQNNLGVFE 622

Qy        635 ECGLWVLTDANL-TKD------------------------------------------ 649
Db        623 DAGLALTTSTNLNTKQRSAAKCPQPANRRRRSSVLLLDSNASKAAEFQDQDLRKCCEDVM 682

Qy        650 ------------------------------YIDGVYD-NAEYAERFM--EENE-GHIVD 674
Db        683 HENPMGYTCEKRAKYIQEGDACKAAFLECCRYIKGVRDENQRESELFLARDDNEDGFIAD 742

Qy        675 IHDFSLGSSPHVRKHFPETWIWL------DTNMGYRIYQEFEVTVPDSITSWVATGFVIS 728
```

```
Db        743 SDIIS-------RSDFPKWWLWLTKDLTEEPNSQGISSKTMSFYLRDSITTWVVLAVSFT 795

Qy        729 EDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSD 788

Db        796 PTKGI-CVAEPYEIRVMKVFFIDLQMPYSVVKNEQVEIRAILHNYVNEDIYVRVELLYNP 854

Qy        789 KFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLG--EIPITVTALSPTASDAVT 846

Db        855 AF----CSASTKGQRYRQQPPIKALSSRAVPFVIVPLEQGLHDVEIKASVQEALWSDGVR 910

Qy        847 QMILVKAEGIEKSYSQSILLDLTDNRLQST------LKTLSFSFPPNTVTGSERVQITAI 900

Db        911 KKLKVVPEGVQKSIVTIVKLDPRAKGVGGTQLEVIKARKLDDRVPDTEIETKIIIQGDPV 970

Qy        901 GDVLGPSING--LASLIRMPYGCGEQNMINFAPNI---YILDYLTKKKQLTDNLKEKALS 955

Db        971 AQIIENSIDGSKLNHLIITPSGCGEQNMIRMAAPVIATYYLDTTEQWETLGINRRTEAVN 1030

Qy        956 FMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF-LEADPYIDIDQNVLHRTY 1014

Db       1031 QIVTGYAQQMVYKKADHSYAAFTN--RASSSWLTAYVVKVFAMAAKMVAGISHEIICGGV 1088

Qy       1015 TWL-KGHQKSNGEFWDPGRVIHSELQG---GNKSPVTLTAYIVTSLLGYR----KYQPNI 1066

Db       1089 RWLILNRQQPDGAFKENAPVLSGTMQGGIQGAEEEVYLTAFILVALLESKTICNDYVNSL 1148

Qy       1067 D--VQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFW 1124

Db       1149 DSSIKKATNYLLKKYEK-LQRPYTTALTAYALA------AADQLNDDRVLMAASTGRDHW 1201

Qy       1125 VSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDT 1184

Db       1202 --------EEYNAHTHNIEGTSYALLALLKMKKFDQTGPIVRWLTDQNFYGETYGQTQAT 1253

Qy       1185 TVALKALSEFAALMNTER-TNIQVTVTGPSSPSPL------------AVVQP---MAVNI 1228

Db       1254 VMAFQALAEYEIQMPTHKDLNLDITIELPDREVPIRYRINYENALLARTVETKLNQDITV 1313

Qy       1229 SANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVD------L 1282

Db       1314 TASGDGKATMTILTFYNAQL----QEKANVCNK--FHLNVSVE--NIHLNAMGAKGALML 1365

Qy       1283 NVCTSFSGPGRSGMALMEVNLLSGFMVPSEAIS-----LSETVKKVEYDHGKLN-----L 1332

Db       1366 KICTRYLGEVDSTMTIIDISMLTGFLPDAEDLTRLSKGVDRYISRYEVDNNMAQKVAVII 1425

Qy       1333 YLDSVNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCS 1391

Db       1426 YLNKVSHSEDECLHPKILKHFEVGFIQPGSVKVYSYYNLDEKCTKFYHPDKGTGLLNKIC 1485

Qy       1392 DVQGCRPCEDGASGSHH--------------HSSVIFIFCFKLLYFME 1425

Db       1486 IGNVCRCAGETCSSLNHQERIDVPLQIEKACETNVDYVYKTKLLRIEE 1533
```

RESULT 11

```
US-09-017-947-35
; Sequence 35, Application US/09017947
; Patent No. 6303754
;  GENERAL INFORMATION:
;    APPLICANT:  VOGEL, CARL-WILHELM
;    APPLICANT:  BREDEHORST, REINHORST
;    APPLICANT:  KOCK, MICHAEL
;    APPLICANT:  FRITZINGER, DAVID
;    TITLE OF INVENTION:  RECOMBINANT PROCVF
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/017,947
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/662,227
;      FILING DATE:  14-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-0107-0X
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-413-3000
;      TELEFAX:  703-413-2220
;  INFORMATION FOR SEQ ID NO:  35:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1648 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-017-947-35

  Query Match              12.0%;  Score 880;  DB 4;  Length 1648;
  Best Local Similarity    22.7%;  Pred. No. 6.6e-64;
  Matches  365;  Conservative  295;  Mismatches  664;  Indels  284;  Gaps  59;

Qy         18 AALAV-----APGPRFLVTAPGIIRPGGNVTIGVELL-EHCPSQVTVKA----ELLKTAS 67

Db         10 AALLIGFPGSSHGALYTLITPAVLRTDTEEQILVEAHGDSTPKQLDIFVHDFPRKQKTLF 69

Qy         68 NLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETK 127

Db         70 QTRVDMNPAGGMLVT---PTIEIPAKEVSTDSRQNQYVVVQVTGPQVRLEKVVLLSYQSS 126

Qy        128 RISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILI--KDPKSNLIQQWLSQQSD 185

Db        127 --FLFIQTDKGIYTPGSPVLYRVFSM--DHHTSKMNKTVIVEFQTPEGILVS---SNSVD 179

Qy        186 LGVISKTFQLSSHPILGDWSI--QVQVNDQTYYQSFQVSEYVLPKFEVTLQTP---LYCS 240

Db        180 LNFF-WPYNLPDLVSLGTWRIVAKYEHSPENYTAYFDVRKYVLPSFEVRLQPSEKFFYID 238

Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNI----TKTFKINGSANFSFND 296

Db        239 GN-ENFHVSITARYLYGEEVEG-VAFVLFGVKIDDAKKSIPDSLTRIPIIDGDGKATLKR 296

Qy        297 EEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDY 356

Db        297 DTFRSRFPNLNELVGHTLYAS------VTVMTESGSDMVVTEQSGIHIVASPYQIHFTKT 350

Qy        357 TTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQ 416

Db        351 PKYFKPGMPYELTVYVTNPDGSPA--------AHVPVVSEAF------HSMGTTLSDGTA 396

Qy        417 KINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDEN 476

Db        397 KLILNIPLNAQ---SLPITVRTNHGDLPRERQATKSMTAIAYQTQGGSGNYLHVAITSTE 453

Qy        477 IKVGS--PFELVVSGN----KRLKELSYMVVSRGQLVAVGKQNS-------TM-FSLTPE 522

Db        454 IKPGDNLPVNFNVKGNANSLKQIKYFTYLILNKGKIFKVGRQPRRDGQNLVTMNLHITPD 513

Qy        523 NSWTPKACVIVYYIEDDGEIISDVLKIPVQ------LVFK--NKIKLYWSKVKAEPSEKV 574

Db        514 LIPSPR--FVAYYQVGNNEIVADSVWVDVKDTCMGTLVVKGDNLIQMPGAAMKIKLEGDF 571
```

```
Qy     575 SLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQ 634
Db     572 GAR---------VGLVAVDKAVYVLNDKYKISQAKIWDTIEKSDFGCTAGSGQNNLGVFE 622

Qy     635 ECGLWVLTDANL-TKD-------------------------------------------- 649
Db     623 DAGLALTTSTNLNTKQRSAAKCPQPANRRRRSSVLLLDSNASKAAEFQDQDLRKCCEDVM 682

Qy     650 ------------------------------YIDGVYD-NAEYAERFM--EENE-GHIVD 674
Db     683 HENPMGYTCEKRAKYIQEGDACKAAFLECCRYIKGVRDENQRESELFLARDDNEDGFIAD 742

Qy     675 IHDFSLGSSPHVRKHFPETWIWL------DTNMGYRIYQEFEVTVPDSITSWVATGFVIS 728
Db     743 SDIIS-------RSDFPKWWLWLTKDLTEEPNSQGISSKTMSFYLRDSITTWVVLAVSFT 795

Qy     729 EDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSD 788
Db     796 PTKGI-CVAEPYEIRVMKVFFIDLQMPYSVVKNEQVEIRAILHNYVNEDIYVRVELLYNP 854

Qy     789 KFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLG--EIPITVTALSPTASDAVT 846
Db     855 AF----CSASTKGQRYRQQPPIKALSSRAVPFVIVPLEQGLHDVEIKASVQEALWSDGVR 910

Qy     847 QMILVKAEGIEKSYSQSILLDLTDNRLQST------LKTLSFSFPPNTVTGSERVQITAI 900
Db     911 KKLKVVPEGVQKSIVTIVKLDPRAKGVGGTQLEVIKARKLDDRVPDTEIETKIIIQGDPV 970

Qy     901 GDVLGPSING--LASLIRMPYGCGEQNMINFAPNI---YILDYLTKKKQLTDNLKEKALS 955
Db     971 AQIIENSIDGSKLNHLIITPSGCGEQNMIRMAAPVIATYYLDTTEQWETLGINRRTEAVN 1030

Qy     956 FMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF-LEADPYIDIDQNVLHRTY 1014
Db    1031 QIVTGYAQQMVYKKADHSYAAFTN--RASSSWLTAYVVKVFAMAAKMVAGISHEIICGGV 1088

Qy    1015 TWL-KGHQKSNGEFWDPGRVIHSELQG---GNKSPVTLTAYIVTSLLGYR----KYQPNI 1066
Db    1089 RWLILNRQQPDGAFKENAPVLSGTMQGGIQGAEEEVYLTAFILVALLESKTICNDYVNSL 1148

Qy    1067 D--VQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFW 1124
Db    1149 DSSIKKATNYLLKKYEK-LQRPYTTALTAYALA------AADQLNDDRVLMAASTGRDHW 1201

Qy    1125 VSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDT 1184
Db    1202 --------EEYNAHTHNIEGTSYALLALLKMKKFDQTGPIVRWLTDQNFYGETYGQTQAT 1253

Qy    1185 TVALKALSEFAALMNTER-TNIQVTVTGPSSPSPL------------AVVQP---MAVNI 1228
Db    1254 VMAFQALAEYEIQMPTHKDLNLDITIELPDREVPIRYRINYENALLARTVETKLNQDITV 1313

Qy    1229 SANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVD------L 1282
Db    1314 TASGDGKATMTILTFYNAQL----QEKANVCNK--FHLNVSVE--NIHLNAMGAKGALML 1365

Qy    1283 NVCTSFSGPGRSGMALMEVNLLSGFMVPSEAIS-----LSETVKKVEYDHGKLN-----L 1332
Db    1366 KICTRYLGEVDSTMTIIDISMLTGFLPDAEDLTRLSKGVDRYISRYEVDNNMAQKVAVII 1425

Qy    1333 YLDSVNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCS 1391
Db    1426 YLNKVSHSEDECLHPKILKHFEVGFIQPGSVKVYSYYNLDEKCTKFYHPDKGTGLLNKIC 1485

Qy    1392 DVQGCRPCEDGASGSHH--------------HSSVIFIFCFKLLYFME 1425
Db    1486 IGNVCRCAGETCSSLNHQERIDVPLQIEKACETNVDYVYKTKLLRIEE 1533
```

RESULT 12
US-08-447-411-76
; Sequence 76, Application US/08447411

```
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  76:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1333 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-447-411-76

  Query Match             10.5%;  Score 768.5;  DB 1;  Length 1333;
  Best Local Similarity   24.4%;  Pred. No. 9.7e-55;
  Matches  301;  Conservative  208;  Mismatches  477;  Indels  247;  Gaps   45;

Qy     349 YIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSG 408
Db      22 YQIYFTKTPKYFKPGMPYELTVYVTKPDGSPAA-------HVPVVSEAIHSE-------G 67

Qy     409 NQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYI 468
Db      68 TTLSDGTAKLFLNTPQNAQ---SLPITVRTNHGDLPRERQAIKSMTATAYQTQGGSGNYL 124

Qy     469 QLKTRDENIKVGS--PFELVVSGN----KRLKELSYMVVSRGQLVAVGKQNS-------T 515
Db     125 HVAITSTEIKPGDNLPVNFNVRGNANSLNQIKYFTYLILNKGKIFKVGRQHRGDGENLVT 184

Qy     516 M-FSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQ------LVFK------NKIKLY 562
Db     185 MNLHITPD--LIPSFRFVAYYQVGNNEIVADSVWVDVKDTCMGTLVVKGATSRDNRIQM- 241

Qy     563 WSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYY 622
Db     242 -------PGAAMKIKLE-GDPGAWIGLVAVDKAEYVLNDKYKISQAKIWDTIEKSDFGCT 293

Qy     623 LGMFMNSFAVFQECGLWVLTDANLT----------------------------------- 647
Db     294 AGSGQNNLGVFEDAGLALTTSTNLNTKQRSAAKCPQPANRRRRSSVLLLDSNASKAAQFQ 353

Qy     648 -----KDYIDGVYDN-----AEYAERFMEENEG---------HIV--------------- 673
```

```
Db        354 DQDLRKCCEDGMHENPMGHTCEKREKYIQEGDACKAAFLECCHYIKGIQDDNKRESELFL 413

Qy        674 ---DIHDFSLGSSPHV-RKHFPETWIWL---------DTNMGYRIYQEFEVTVPDSITSW 720
                 |  |   |         |  ||:|:||              :  :|    :    : ||||:|
Db        414 ARSDFEDDLFGEGNITSRSDFPESWLWLMEQLSEHPNSKGISSKIVPFY---LRDSITTW 470

Qy        721 VATGFVISEDLGLGLTTT-------PVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNY 773
                      | | :||: |        | |:    :  |||  | |||||:: |:   :    ::||
Db        471 --------ELLAVGLSPTKGICVAEPYEITVMKDFFIDLQLPYSVVKNEQVKIRAVLYNY 522

Qy        774 LKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLG--EIP 831
                  |:| :   |   |         | |:       ::: | :  :         | | | |   |    ::
Db        523 ADKDIYVRVELLYSPAFCSASTESQ----RYREQLPIKALSSRAVSFVIVPLEQGLHDVE 578

Qy        832 ITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQST------LKTLSFSFP 885
              :| :        || | : : |  ||  ||     | ||          :  |           |:   |
Db        579 VTASVQGELMSDGVKKKLKVVPEGEWKSIVTIIELDPHTKGIGGTQVELVKANKLNDRVP 638

Qy        886 PNTVTGSERVQITAIGDVLGPSING--LASLIRMPYGCGEQNMINF-APNI--YILDYLT 940
                   :        :|    :     :   ||:|   |   ||    |:||||||||    || |   | ||
Db        639 DTEIETKITIQGDPVAQTIENSIDGSKLNHLIITPFGCGEQNMIRMTAPVIATYYLDTTQ 698

Qy        941 KKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF-LEA 999
              : :  |    | : :|:: :    || ::|:|:: |  |::|| |     :  |:||:|:|:: | | |
Db        699 QWETLGINRRTEAVNQIMTGYAQQLVYKKADHSYAAFTN--SASSSWLTAYVVKIFALAA 756

Qy       1000 DPYIDIDQNVLHRTYTWL-KGHQKSNGEFWDPGRVIHSELQGGNKSPV---TLTAYIVTS 1055
                ||:  ::         ||       |:::|  |:      |:       :|||  :            :|||:|: :
Db        757 KIVKDINHEIVCGGMRWLILNRQRTDGVFRENAPVLFGTMQGGIQGAEPEGSLTAFILVA 816

Qy       1056 LLGYRKYQPNIDVQESIHFLESEFSRG----------ISDNYTLALITYALSSVGSPKAK 1105
              ||  |            |: |:|  |:                :  || ||  |||:              |
Db        817 LLESRSI-----CNAYINILDSSISKATDYLLKKYEKLQRPYTTALTAYALA------AA 865

Qy       1106 EALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIM 1165
               | ||            |     |        :               : :||   :||||:        :    |   |::
Db        866 ERLNDDRVLMAASTGRNRWEEPNA--------HTHNIEGTSYALLALLKMKKFVEAGPVV 917

Qy       1166 RWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTER-TNIQVTVTGPSSPSP------- 1217
              :||  |:    | :   ||  | :   :||:|:    |  |  :    |: :|:    |     |
Db        918 QWLIDQQYYGGTYGQTQATVMMFQALAEYEIQMPTHKDLNLDITIELPDREVPIRYRINY 977

Qy       1218 ----LAVVQPMAVN----ISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVA 1269
                      ||             :|   :|:|:|  |     |||:    |       ::  |:    |   |||:
Db        978 ENALLAQTVETKLNEDFTVSASGDGKATMTILTVYNAQL----REDANVCNK--FHLDVS 1031

Qy       1270 V-------KENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAIS-LSETVK 1321
              |          || |     | :||:|  |       |  :::|:|:|:||:   :|  ::  ||: |
Db       1032 VENVQLNLKEAKGAKGALKLKICTRYLGEVDSTMTIIDVSMLTGFVPDTEDLTRLSKGVD 1091

Qy       1322 K----VEYDH-----GKLNLYLDSVNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPR 1371
              :      | ::        | : :|||  |: ::  |::     :::|:|     |     ||  :   ||
Db       1092 RYISMFEINNNMAQKGTVIIYLDKVSHSEDECLHFKILKHFEVGFIQPGSVKVYSYYNLD 1151

Qy       1372 RQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGAS 1404
                 :  :  |:  :         :            ||  |:  |
Db       1152 EKCTKIYHPDEATGLLNKICVGNVCRCAEETCS 1184
```

RESULT 13
US-08-662-227-34
; Sequence 34, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
;   APPLICANT: VOGEL, CARL-WILHELM
;   APPLICANT: BREDEHORST, REINHORST
;   APPLICANT: KOCK, MICHAEL
;   APPLICANT: FRITZINGER, DAVID
;   TITLE OF INVENTION: RECOMBINANT PROCVF
;   NUMBER OF SEQUENCES: 39
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/662,227
;   FILING DATE: 14-JUN-1996
;   CLASSIFICATION: 530
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 1126-0107-0X
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 34:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1333 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-662-227-34

```
  Query Match             10.3%;  Score 758.5;  DB 2;  Length 1333;
  Best Local Similarity   24.4%;  Pred. No. 6.6e-54;
  Matches  301;  Conservative  208;  Mismatches  477;  Indels  247;  Gaps   46;

Qy        349 YIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSG 408
                 | | |           || : :   || ||: ||:                 |    | |::  ::|         |
Db         22 YQIYFTKTPKYFKPGMPYELTVYVTKPDGSPAA-------HVPVVSEAIHSE-------G 67

Qy        409 NQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYI 468
                    :    |:     ||:          ||     ::     |          || |      :   |  |:
Db         68 TTLSDGTAKLFLNTPQNAQ---SLPITVRTNHGDLPRERQAIKSMTATAYQTQGGSGNYL 124

Qy        469 QLKTRDENIKVGS--PFELVVSGN----KRLKELSYMVVSRGQLVAVGKQNS------TM 516
                       ||  |      | ||         ::|   :|::::|::   ||:|:          ||
Db        125 HVAITSTEIKPGDNLPVNFNVRGNANSLNQIKYFTYLILNKGKIFKVGRQHRGDGNLVTM 184

Qy        517 -FSLTPENSWTPKACVIVYYIEDDGEI-ISDVLKIPVQ------LVFK------NKIKLY 562
                :||:      |    ||   ::| : : |:    :          |||             |:|::
Db        185 NLHITPD--LIPSFRFVAYYQVGNNEIEVADSVWVDVKDTCMGTLVVKGATSRDNRIQM- 241

Qy        563 WSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYY 622
                        |     : :::       | : :|:|||||:  ::|       |:     :     :|   :  |
Db        242 -------PGAAMKIKLE-GDPGAWIGLVAVDKAEYVLNDKYKISQAKIWDTIEKSDFGCT 293

Qy        623 LGMFMNSFAVFQECGLWVLTDANLT----------------------------------- 647
                 |     |:  ||::  || :  |   ||
Db        294 AGSGQNNLGVFEDAGLALTTSTNLNTKQRSAAKCPQPANRRRRSSVLLLDSNASKAAQFQ 353

Qy        648 -----KDYIDGVYDN-----AEYAERFMEENEG---------HIV--------------- 673
                     |  ||:::|          |  |::::|  :             | :
Db        354 DQDLRKCCEDGMHENPMGHTCEKREKYIQEGDACKAAFLECCHYIKGIQDDNKRESELFL 413

Qy        674 ---DIHDFSLGSSPHV-RKHFPETWIWL---------DTNMGYRIYQEFEVTVPDSITSW 720
                 |  |   |         |  ||:|:||              :  :|    :    : ||||:|
Db        414 ARSDFEDDLFGEGNITSRSDFPESWLWLMEQLSEHPNSKGISSKIVPFY---LRDSITTW 470

Qy        721 VATGFVISEDLGLGLTTT-------PVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNY 773
                      | | :||: |        | |:    :  |||  | |||||:: |:   :    ::||
Db        471 --------ELLAVGLSPTKGICVAEPYEITVMKDFFIDLQLPYSVVKNEQVKIRAVLYNY 522

Qy        774 LKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLG--EIP 831
                  |:| :   |   |         | |:       ::: | :  :         | | | |   |    ::
Db        523 ADKDIYVRVELLYSPAFCSASTESQ----RYREQLPIKALSSRAVSFVIVPLEQGLHDVE 578
```

```
Qy        832 ITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQST------LKTLSFSFP 885
Db        579 VTASVQGELMSDGVKKKLKVVPEGEWKSIVTIIELDPHTKGIGGTQVELVKANKLNDRVP 638

Qy        886 PNTVTGSERVQITAIGDVLGPSING--LASLIRMPYGCGEQNMINF-APNI--YILDYLT 940
Db        639 DTEIETKITIQGDPVAQTIENSIDGSKLNHLIITPFGCGEQNMIRMTAPVIATYYLDTTQ 698

Qy        941 KKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF-LEA 999
Db        699 QWETLGINRRTEAVNQIMTGYAQQLVYKKADHSYAAFTN--SASSSWLTAYVVKIFALAA 756

Qy       1000 DPYIDIDQNVLHRTYTWL-KGHQKSNGEFWDPGRVIHSELQGGNKSPV---TLTAYIVTS 1055
Db        757 KIVKDINHEIVCGGMRWLILNRQRTDGVFRENAPVLFGTMQGGIQGAEPEGSLTAFILVA 816

Qy       1056 LLGYRKYQPNIDVQESIHFLESEFSRG----------ISDNYTLALITYALSSVGSPKAK 1105
Db        817 LLESRSI-----CNAYINILDSSISKATDYLLKKYEKLQRPYTTALTAYALA------AA 865

Qy       1106 EALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIM 1165
Db        866 ERLNDDRVLMAASTGRNRWEEPNA--------HTHNIEGTSYALLALLKMKKFVEAGPVV 917

Qy       1166 RWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTER-TNIQVTVTGPSSPSP------- 1217
Db        918 QWLIDQQYYGGTYGQTQATVMMFQALAEYEIQMPTHKDLNLDITIELPDREVPIRYRINY 977

Qy       1218 ----LAVVQPMAVN----ISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVA 1269
Db        978 ENALLAQTVETKLNEDFTVSASGDGKATMTILTVYNAQL----REDANVCNK--FHLDVS 1031

Qy       1270 V-------KENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAIS-LSETVK 1321
Db       1032 VENVQLNLKEAKGAKGALKLKICTRYLGEVDSTMTIIDVSMLTGFVPDTEDLTRLSKGVD 1091

Qy       1322 K----VEYDH-----GKLNLYLDSVNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPR 1371
Db       1092 RYISMFEINNNMAQKGTVIIYLDKVSHSEDECLHFKILKHFEVGFIQPGSVKVYSYYNLD 1151

Qy       1372 RQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGAS 1404
Db       1152 EKCTKIYHPDEATGLLNKICVGNVCRCAEETCS 1184
```

RESULT 14
US-09-017-947-34
; Sequence 34, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
;   APPLICANT: VOGEL, CARL-WILHELM
;   APPLICANT: BREDEHORST, REINHORST
;   APPLICANT: KOCK, MICHAEL
;   APPLICANT: FRITZINGER, DAVID
;   TITLE OF INVENTION: RECOMBINANT PROCVF
;   NUMBER OF SEQUENCES: 39
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
;   CITY: ARLINGTON
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/017,947
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/662,227
;   FILING DATE: 14-JUN-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: OBLON, NORMAN F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 1126-0107-0X
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 703-413-3000
;   TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 34:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1333 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-017-947-34

```
  Query Match             10.3%;  Score 758.5;  DB 4;  Length 1333;
  Best Local Similarity   24.4%;  Pred. No. 6.6e-54;
  Matches  301;  Conservative  208;  Mismatches  477;  Indels  247;  Gaps   46;

Qy        349 YIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSG 408
Db         22 YQIYFTKTPKYFKPGMPYELTVYVTKPDGSPAA------HVPVVSEAIHSE-------G 67

Qy        409 NQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYI 468
Db         68 TTLSDGTAKLFLNTPQNAQ---SLPITVRTNHGDLPRERQAIKSMTATAYQTQGGSGNYL 124

Qy        469 QLKTRDENIKVGS--PFELVVSGN----KRLKELSYMVVSRGQLVAVGKQNS------TM 516
Db        125 HVAITSTEIKPGDNLPVNFNVRGNANSLNQIKYFTYLILNKGKIFKVGRQHRGDGNLVTM 184

Qy        517 -FSLTPENSWTPKACVIVYYIEDDGEI-ISDVLKIPVQ------LVFK------NKIKLY 562
Db        185 NLHITPD--LIPSFRFVAYYQVGNNEIEVADSVWVDVKDTCMGTLVVKGATSRDNRIQM- 241

Qy        563 WSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYY 622
Db        242 -------PGAAMKIKLE-GDPGAWIGLVAVDKAEYVLNDKYKISQAKIWDTIEKSDFGCT 293

Qy        623 LGMFMNSFAVFQECGLWVLTDANLT----------------------------------- 647
Db        294 AGSGQNNLGVFEDAGLALTTSTNLNTKQRSAAKCPQPANRRRRSSVLLLDSNASKAAQFQ 353

Qy        648 -----KDYIDGVYDN-----AEYAERFMEENEG---------HIV-------------- 673
Db        354 DQDLRKCCEDGMHENPMGHTCEKREKYIQEGDACKAAFLECCHYIKGIQDDNKRESELFL 413

Qy        674 ---DIHDFSLGSSPHV-RKHFPETWIWL---------DTNMGYRIYQEFEVTVPDSITSW 720
Db        414 ARSDFEDDLFGEGNITSRSDFPESWLWLMEQLSEHPNSKGISSKIVPFY---LRDSITTW 470

Qy        721 VATGFVISEDLGLGLTTT-------PVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNY 773
Db        471 --------ELLAVGLSPTKGICVAEPYEITVMKDFFIDLQLPYSVVKNEQVKIRAVLYNY 522

Qy        774 LKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLG--EIP 831
Db        523 ADKDIYVRVELLYSPAFCSASTESQ----RYREQLPIKALSSRAVSFVIVPLEQGLHDVE 578

Qy        832 ITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQST------LKTLSFSFP 885
Db        579 VTASVQGELMSDGVKKKLKVVPEGEWKSIVTIIELDPHTKGIGGTQVELVKANKLNDRVP 638

Qy        886 PNTVTGSERVQITAIGDVLGPSING--LASLIRMPYGCGEQNMINF-APNI--YILDYLT 940
Db        639 DTEIETKITIQGDPVAQTIENSIDGSKLNHLIITPFGCGEQNMIRMTAPVIATYYLDTTQ 698

Qy        941 KKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF-LEA 999
```

```
Db        699 QWETLGINRRTEAVNQIMTGYAQQLVYKKADHSYAAFTN--SASSSWLTAYVVKIFALAA 756

Qy       1000 DPYIDIDQNVLHRTYTWL-KGHQKSNGEFWDPGRVIHSELQGGNKSPV---TLTAYIVTS 1055

Db        757 KIVKDINHEIVCGGMRWLILNRQRTDGVFRENAPVLFGTMQGGIQGAEPEGSLTAFILVA 816

Qy       1056 LLGYRKYQPNIDVQESIHFLESEFSRG----------ISDNYTLALITYALSSVGSPKAK 1105

Db        817 LLESRSI-----CNAYINILDSSISKATDYLLKKYEKLQRPYTTALTAYALA------AA 865

Qy       1106 EALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQTSEGIPIM 1165

Db        866 ERLNDDRVLMAASTGRNRWEEPNA--------HTHNIEGTSYALLALLKMKKFVEAGPVV 917

Qy       1166 RWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTER-TNIQVTVTGPSSPSP------- 1217

Db        918 QWLIDQQYYGGTYGQTQATVMMFQALAEYEIQMPTHKDLNLDITIELPDREVPIRYRINY 977

Qy       1218 ----LAVVQPMAVN----ISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVA 1269

Db        978 ENALLAQTVETKLNEDFTVSASGDGKATMTILTVYNAQL----REDANVCNK--FHLDVS 1031

Qy       1270 V-------KENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAIS-LSETVK 1321

Db       1032 VENVQLNLKEAKGAKGALKLKICTRYLGEVDSTMTIIDVSMLTGFVPDTEDLTRLSKGVD 1091

Qy       1322 K----VEYDH-----GKLNLYLDSVNETQ-FCVNIPAVRNFKVSNTQDASVSIVDYYEPR 1371

Db       1092 RYISMFEINNNMAQKGTVIIYLDKVSHSEDECLHFKILKHFEVGFIQPGSVKVYSYYNLD 1151

Qy       1372 RQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGAS 1404

Db       1152 EKCTKIYHPDEATGLLNKICVGNVCRCAEETCS 1184
```

RESULT 15
US-07-779-172A-3
; Sequence 3, Application US/07779172A
; Patent No. 5322838
; GENERAL INFORMATION:
; APPLICANT: Silberstein, David S.
; APPLICANT: Minkoff, Marjorie
; TITLE OF INVENTION: INHIB, A Factor Which Inhibits
; TITLE OF INVENTION: Cytokine-Activated Leukocytes
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-2678
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/779,172A
; FILING DATE: 19911016
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fox, Samuel L.
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0627.1270001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; TELEX: 248636 SSK
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 645 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: both
; MOLECULE TYPE: peptide
US-07-779-172A-3

Query Match 4.5%; Score 329; DB 1; Length 645;
Best Local Similarity 20.8%; Pred. No. 1.6e-18;
Matches 143; Conservative 133; Mismatches 270; Indels 142; Gaps 22;

```
Qy         26 PRFLVTAPGIIR-------------PGGNVTIGVELLEH------CPSQVTVKAELLKTA 66

Db          2 PMYSIITPNILRLESEETMVLEAHDAQGDVPVTVTVHDFPGKKLVLSSEKTVLTPATNHM 61

Qy         67 SNLTVSV-LEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQ--DEILFSNSTRLS 123

Db         62 GNVTFTIPANREFKSEKGRNKFVT--------------VQATFBTQVVEKVLL-----VS 102

Qy        124 FETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQ-WLSQ 182

Db        103 LQSGYL--FIQTDKTIYTPGSTVLYRIFTVNHKLLPVBRTVMVNIENPEGIPVKQDSLSS 160

Qy        183 QSDLGVISKTFQLSSHPILGDWSIQVQVND---QTYYQSFQVSEYVLPKFEVTLQ--TPL 237

Db        161 QNQLGVLPLSWDIPELVNMGQWKIRAYYENSPQQVFSTEFEVKEYVLPSFEVIVEPTEKF 220

Qy        238 YCSMNSKHLNGTITAKYTYGKPVKGDVTLTF-----------------LPLSFWGKKKNI 280

Db        221 YYIYNEKGLEVTITARFLYGKKVEGTAFVIFGIQDGEQRISLPESLKRIPIEDGSGEVVL 280

Qy        281 TKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVST 340

Db        281 SRKVLLDGVQNLRAEDLVGKSLYVSA------------------TVILHSGSDMVQAERS 322

Qy        341 NVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGN---QLTLEERRNNVVITVTQRN 397

Db        323 GIPIVTSPYQIHFTKTPKYFKPGMPFDLMVFVTNPDGSPAYRVPVAVQGEDTVQSLTQGD 382

Qy        398 YTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVH 457

Db        383 GVAKLSINTHPSQK-----------PLSITVRTKKQELSEAEQ--------ATRTMQALP 423

Qy        458 SLFKSPSKTYIQLKTRDENIKVGSPFELVV------SGNKRLKELSYMVVSRGQLVAVGK 511

Db        424 YSTVGNSNNYLHLSVLTTELRPGETLNVNFLLRMDRAHEAKIRYYTYLIMNKGRLLKAGR 483

Qy        512 ------QNSTMFSLTPENSWTPKACVIVYY----IEDDGEIISDVLKIPVQ------LVF 555

Db        484 QVREPGQDLVVLPLSITTDFIPSFRLVAYYTVIGASGQREVVADSVWVDVKDSCVGSLVV 543

Qy        556 KNKIKLYWSKVKAE-----PSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENV 610

Db        544 KSG--------QSEDRQPVPGQQMTLKIEGDHGARVV-LVAVDKGVFVLNKKNKLTQSKI 594

Qy        611 VHELELYNTGYYLGMFMNSFAVFQECGL 638

Db        595 WDVVEKADIGCTPGSGKDYAGVFSDAGL 622
```

Search completed: January 15, 2004, 18:15:19
Job time : 33 secs

```
OM protein - protein search, using sw model

Run on:          January 15, 2004, 18:13:05 ; Search time 44 Seconds
                                             (without alignments)
                                             6636.344 Million cell updates/sec

Title:           US-10-020-095-4
Perfect score:   7348
Sequence:        1 MQGPPLLTAAHLLCVCTAAL..........HSSVIFIFCFKLLYFMELWL 1428

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters:      762491

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
                 1:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
                 2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
                 3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
                 4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
                 5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
                 6:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
                 7:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
                 8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
                 9:  /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
                 10:  /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
                 11:  /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
                 12:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
                 13:  /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
                 14:  /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
                 15:  /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
                 16:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
                 17:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
                 18:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3360	45.7	665	12	US-10-108-260A-3396	Sequence 3396, Ap
2	1854.5	25.2	1508	12	US-10-369-493-5314	Sequence 5314, Ap
3	1850	25.2	1519	12	US-10-369-493-5313	Sequence 5313, Ap
4	1441	19.6	1474	12	US-10-292-081A-15	Sequence 15, Appl
5	1441	19.6	1474	14	US-10-052-817-2	Sequence 2, Appli
6	1440	19.6	1474	9	US-09-873-403-5	Sequence 5, Appli
7	1440	19.6	1474	12	US-10-292-081A-10	Sequence 10, Appl
8	1440	19.6	1474	12	US-10-292-081A-12	Sequence 12, Appl
9	1440	19.6	1474	12	US-10-292-081A-13	Sequence 13, Appl
10	1440	19.6	1474	12	US-10-331-496A-38	Sequence 38, Appl
11	1437.5	19.6	1508	12	US-09-756-247-4	Sequence 4, Appli
12	1433	19.5	1450	12	US-09-756-247-23	Sequence 23, Appl
13	1433	19.5	1491	12	US-09-756-247-25	Sequence 25, Appl
14	1428	19.4	1500	12	US-10-292-081A-9	Sequence 9, Appli
15	1421.5	19.3	1492	12	US-09-981-151A-10	Sequence 10, Appl
16	1419	19.3	1451	12	US-09-756-247-24	Sequence 24, Appl
17	1384	18.8	1500	12	US-10-316-253-267	Sequence 267, App
18	1384	18.8	1500	12	US-10-316-253-269	Sequence 269, App
19	1255	17.1	1285	9	US-09-925-301-1394	Sequence 1394, Ap
20	1157	15.7	1411	12	US-10-094-886-124	Sequence 124, App
21	1113.5	15.2	941	12	US-09-981-151A-39	Sequence 39, Appl
22	1109.5	15.1	1436	12	US-10-094-886-126	Sequence 126, App
23	1086	14.8	936	12	US-09-981-151A-35	Sequence 35, Appl
24	1080	14.7	898	12	US-09-981-151A-36	Sequence 36, Appl
25	1075	14.6	936	12	US-09-981-151A-37	Sequence 37, Appl
26	1059.5	14.4	931	12	US-09-981-151A-38	Sequence 38, Appl
27	960	13.1	1661	11	US-09-842-758-42	Sequence 42, Appl
28	956	13.0	1663	11	US-09-842-758-43	Sequence 43, Appl
29	934	12.7	1663	11	US-09-842-758-41	Sequence 41, Appl
30	931	12.7	1663	9	US-09-875-519A-22	Sequence 22, Appl
31	921	12.5	1602	9	US-09-778-927A-59	Sequence 59, Appl
32	894	12.2	1642	10	US-09-925-442-2	Sequence 2, Appli
33	880	12.0	1648	10	US-09-925-442-35	Sequence 35, Appl
34	804.5	10.9	751	12	US-09-981-151A-80	Sequence 80, Appl
35	796	10.8	1612	11	US-09-842-758-6	Sequence 6, Appli
36	758.5	10.3	1333	10	US-09-925-442-34	Sequence 34, Appl
37	748	10.2	1251	9	US-09-778-927A-58	Sequence 58, Appl
38	747	10.2	643	12	US-10-292-081A-11	Sequence 11, Appl
39	747	10.2	643	12	US-10-087-188-2	Sequence 2, Appli
40	738	10.0	912	12	US-09-756-247-29	Sequence 29, Appl
41	733.5	10.0	882	12	US-09-756-247-39	Sequence 39, Appl
42	719.5	9.8	528	11	US-09-764-876-20	Sequence 20, Appl
43	718.5	9.8	531	9	US-09-764-853-583	Sequence 583, App
44	689	9.4	134	12	US-10-160-162-139	Sequence 139, App
45	689	9.4	134	12	US-09-820-649-139	Sequence 139, App

ALIGNMENTS

```
RESULT 1
US-10-108-260A-3396
; Sequence 3396, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20040005560A1el full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/10/108,260A
;  CURRENT FILING DATE:  2002-03-27
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3396
;   LENGTH: 665
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-108-260A-3396

  Query Match             45.7%;  Score 3360;  DB 12;  Length 665;
  Best Local Similarity   100.0%;  Pred. No. 4.9e-268;
  Matches  657;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKA 60

Qy         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNST 120

Qy        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWL 180

Qy        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 SQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCS 240
```

```
Qy        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 MNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMK 300

Qy        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 NVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVL 360

Qy        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 KPSLNFTATVKVTRADGNQLTLEERRNNVVITVTQRNYTEYWSGSNSGNQKMEAVQKINY 420

Qy        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

Qy        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 SPFELVVSGNKRLKELSYMVVSRGQLVAVGKQNSTMFSLTPENSWTPKACVIVYYIEDDG 540

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600

Qy        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDN 657
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ASNDITMENVVHELELYNTGYYLGMFMNSFAVFQECGLWVLTDANLTKDYIDGVYDN 657


RESULT 2
US-10-369-493-5314
; Sequence 5314, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5314
;   LENGTH: 1508
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-10-369-493-5314

  Query Match             25.2%;  Score 1854.5;  DB 12;  Length 1508;
  Best Local Similarity   31.8%;  Pred. No. 4.2e-143;
  Matches  476;  Conservative  292;  Mismatches  562;  Indels  169;  Gaps   44;

Qy         15 VCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVL 74
Db         28 VSTTAAPVKPA-TYMLVAPAVVRPDQPFSVCMNLLKQATDEDMIVRIEVRTERNETIAAR 86

Qy         75 EAEGVFEKGSFKTLTLPSLPLNSAD--EIYELRVTGRTQD-EILFSNSTRLSFETKRISV 131
Db         87 VISNL-KPGIAQTVSLSEMPAQSLTPRQSYKLYIRGETLNAELIFENENELKYDQKALSV 145

Qy        132 FIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISK 191
Db        146 FIQTDRAIYRPASLVRYRAIVVKSDLKPYVGNATIKIFDPSRNLISQTIGVTLDRGVYSG 205

Qy        192 TFQLSSHPILGDWSIQVQ----VNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLN 247
Db        206 ELQLAEETLLGDWFIEVETSNGVQDKS---SFTVDTYVLPKFEVNIKTSSFITIND-DLS 261

Qy        248 GTITAKYTYGKPVKGDVTLTF-LPLSFW-----------GKKKN----ITKTFKIN--GS 289
Db        262 VFVDAKYTYGKGVAGKAKVSLELPWHRWHAMVPTIIDENGVKKEEELMVERTVKLNRQGE 321

Qy        290 ANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDY 349
Db        322 AAVVFSNDELKR--------HKLLHEWGGGSIRIVASVTEDITEIERNATHQISTFREEV 373

Qy        350 IIEFFDYTTVLKPSLNFTATVKVTRADGN--QLTLEERRNNVVITVTQRNYTEYWSGSNS 407
Db        374 KLDVEKQGDTFKPGLTYNVVVALKQMDDTPVKATLPKR----VQVSTFYNYP--YNHDTS 427

Qy        408 GNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK------ 461
Db        428 SLQEEKETKIVEVDAHGTSVLTLQPPINCTSARIEAH-YDIGGKDNFTATPIYSSLYVEA 486

Qy        462 --SPSKTYIQLKTRDEN-IKVGSPFELVVSGNKRLKELSYMVVSRGQLV-----AVGKQN 513
Db        487 AVSPTKSFLQLLADNEGAVDVGKSLSFSLKATQPLSTITYQVMSRSNIVVSQQMTVNSEH 546

Qy        514 STMFSLTPENSWTPKACVIVY-YIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSE 572
Db        547 ATI-SFPATANMAPKSRLIVYAIIESSQEVLVDALDFKVEGIFQNQVALSIDKQAVEPGQ 605

Qy        573 KVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMF------ 626
Db        606 NVKFKVT-SDKNSFVGLLVVDQSVLLLKTGNDITREKVEQDLENYDSNNVGGGFGGPRPW 664

Qy        627 ------------------MNSFAVFQECGLWVLTDANLTKD---------YIDGVYDNA 658
Db        665 EAIDRKKRSIWRPWWGIGGSDAQSIFSNAGLVVLTDALLYREPQREFMSVMMMDGAPGMA 724

Qy        659 EYAERFMEENEGHIVDIHDFSLGSS---PHVRKHFPETWIWLDTNMGYRIYQEFEVTVPD 715
Db        725 EAA--FAAPPMG----------GSSPPPPTVRKFFPHTWIWSDLN-STSGEVEMEIEAPD 771

Qy        716 SITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLK 775
Db        772 TITSWVASTFAINEENGLGVAPTTSKLRVFRPFFIQLNLPYAVRRGEKFALLVLVFNYME 831

Qy        776 DATEVKVII--EKSDKFDILMTSSEI---NATGHQQTLLVPSEDGAT---VLFPIRPTHL 827
Db        832 KEQDVTVTLKYDKDSGYDLLKKDGTVVRRDEVGQQNVRIVSVAGGGTSKAVYFPIVPSSI 891

Qy        828 GEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPN 887
Db        892 GEIPVHISAIASQGGDAVEMNLRVDPQGYKVDRNIPFVIDLNNNSSDFS-KNLELIWPND 950

Qy        888 TVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTD 947
Db        951 VVDGSQKARLDVIGDMMGPVLNNAHKLVQMPYGCGEQNMLNLVPNILVVKYLRATNRNES 1010

Qy        948 NLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQ 1007
Db       1011 QLETKAIKFIEQGIQRELTYKRADNSFSAFGDSDKAGSTWLTAFVVRSFHHAKQYAFVDP 1070

Qy       1008 NVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG-NKSPVTLTAYIVTSLLGYRKYQPNI 1066
Db       1071 NVISRAVAFLNSQQMESGAFAERGEVHHKDMQGGAQDGGVALTAFVLISIL-----ENGM 1125

Qy       1067 DVQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWR--AEQEGGMQFW 1124
Db       1126 ENGKAVTYLEKHLDEVSGNAYTMAVVAYALQLAKSKQAGKAFENLKKHKIVEKSGDVKFA 1185

Qy       1125 VSSE--SKLSDS----WQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRWLSRQRNSLGGF 1178
Db       1186 SAQKKVEKLKESRAYMFQARPVDIETTSYAVLSYLAQNQTSESLSIIRWLVSQRNELGGF 1245

Qy       1179 ASTQDTTVALKALSEFAALMNTERTNIQVTV-TGPSSPS------PLAVVQPM------- 1224
Db       1246 TSTQDTVMALQALSSYAAVTYSDKHTSQVTILNGKHTHSFDINIRNAIVLQSYQLSSLND 1305

Qy       1225 AVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNV 1284
```

```
Db       1306 AVSINANGTGVVFAQLSYSY---------YRDSLNDDAPFFCSQEIKEIRAG-NRLQLDL 1355

Qy       1285 CTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVK-----KVEYDHGKLNLYLDSVNE 1339

Db       1356 CCNYTRPGKSNMALAEIDALSGYRFDAEQVHTLTSIEDLQRVEMEKDDTKMNVYFNPLGG 1415

Qy       1340 TQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRP 1398

Db       1416 RPVCLSLYSDVTYQVADQKPANFRLVDYYDPEEQLKMTYAAKQTRSLQEKCG--EDCWP 1472


RESULT 3
US-10-369-493-5313
; Sequence 5313, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
;  APPLICANT: Cao, Yongwei
;  APPLICANT:  Hinkle, Gregory J.
;  APPLICANT:  Slater, Steven C.
;  APPLICANT:  Goldman, Barry S.
;  APPLICANT:  Chen, Xianfeng
;  TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;  TITLE OF INVENTION:  PLANTS WITH IMPROVED PROPERTIES
;  FILE REFERENCE: 38-10(52052)B
;  CURRENT APPLICATION NUMBER: US/10/369,493
;  CURRENT FILING DATE:  2003-02-28
;  PRIOR APPLICATION NUMBER: US 60/360,039
;  PRIOR FILING DATE: 2002-02-21
;  NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5313
;   LENGTH: 1519
;   TYPE: PRT
;   ORGANISM: Caenorhabditis elegans
US-10-369-493-5313

  Query Match             25.2%;  Score 1850;  DB 12;  Length 1519;
  Best Local Similarity   31.6%;  Pred. No. 9.9e-143;
  Matches  477;  Conservative  290;  Mismatches  563;  Indels  180;  Gaps   44;

Qy         15 VCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVL 74

Db         28 VSTTAAPVKPA-TYMLVAPAVVRPDQPFSVCMNLLKQATDEDMIVRIEVRTERNETIAAR 86

Qy         75 EAEGVFEKGSFKTLTLPSLPLNSAD--EIYELRVTGRTQD-EILFSNSTRLSFETKRISV 131

Db         87 VISNL-KPGIAQTVSLSEMPAQSLTPRQSYKLYIRGETLNAELIFENENELKYDQKALSV 145

Qy        132 FIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNILIKDPKSNLIQQWLSQQSDLGVISK 191

Db        146 FIQTDRAIYRPASLVRYRAIVVKSDLKPYVGNATIKIFDPSRNLISQTIGVTLDRGVYSG 205

Qy        192 TFQLSSHPILGDWSIQVQ----VNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLN 247

Db        206 ELQLAEETLLGDWFIEVETSNGVQDKS---SFTVDTYVLPKFEVNIKTSSFITIND-DLS 261

Qy        248 GTITAKYTYGKPVKGDVTLTF-LPLSFW-----------GKKKN----ITKTFKIN--GS 289

Db        262 VFVDAKYTYGKGVAGKAKVSLELPWHRWHAMVPTIIDENGVKKEEELMVERTVKLNRQGE 321

Qy        290 ANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGISRNVSTNVFFKQHDY 349

Db        322 AAVVFSNDELKR--------HKLLHEWGGGSIRIVASVTEDITEIERNATHQISTFREEV 373

Qy        350 IIEFFDYTTVLKPSLNFTATVKVTRADGN--QLTLEERRNNVVITVTQRNYTEYWSGSNS 407

Db        374 KLDVEKQGDTFKPGLTYNVVVALKQMDDTPVKATLPKR----VQVSTFYNYP--YNHDTS 427

Qy        408 GNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGSKSSMAVHSLFK------ 461

Db        428 SLQEEKETKIVEVDAHGTSVLTLQPPINCTSARIEAH-YDIGGKDNFTATPIYSSLYVEA 486

Qy        462 --SPSKTYIQLKTRDEN-IKVGSPFELVVSGNKRLKELSYMVVSRGQLV-----AVGKQN 513

Db        487 AVSPTKSFLQLLADNEGAVDVGKSLSFSLKATQPLSTITYQVMSRSNIVVSQQMTVNSEH 546

Qy        514 STMFSLTPENSWTPKACVIVY-YIEDDGEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSE 572

Db        547 ATI-SFPATANMAPKSRLIVYAIIESSQEVLVDALDFKVEGIFQNQVALSIDKQAVEPGQ 605

Qy        573 KVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLGMF------ 626

Db        606 NVKFKVT-SDKNSFVGLLVVDQSVLLLKTGNDITREKVEQDLENYDSNNVGGGFGGPRPW 664

Qy        627 -------------------MNSFAVFQECGLWVLTDA--------------------NLT 647

Db        665 EAIDRKKRSIWRPWWGIGGSDAQSIFSNAGLVVLTDALLYREPQREFMSERRLNTPGGLT 724

Qy        648 KDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGSS---PHVRKHFPETWIWLDTNMGYR 704

Db        725 VMMMDGAPGMAEAA--FAAPPMG----------GSSPPPPTVRKFFPHTWIWSDLN-STS 771

Qy        705 IYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEF 764

Db        772 GEVEMEIEAPDTITSWVASTFAINEENGLGVAPTTSKLRVFRPFFIQLNLPYAVRRGEKF 831

Qy        765 ALEITIFNYLKDATEVKVII--EKSDKFDILMTSSEI---NATGHQQTLLVPSEDGAT-- 817

Db        832 ALLVLVFNYMEKEQDVTVTLKYDKDSGYDLLKKDGTVVRRDEVGQQNVRIVSVAGGGTSK 891

Qy        818 -VLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQST 876

Db        892 AVYFPIVPSSIGEIPVHISAIASQGGDAVEMNLRVDPQGYKVDRNIPFVIDLNNNSSDFS 951

Qy        877 LKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYIL 936

Db        952 -KNLELIWPNDVVDGSQKARLDVIGDMMGPVLNNAHKLVQMPYGCGEQNMLNLVPNILVV 1010

Qy        937 DYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCF 996

Db       1011 KYLRATNRNESQLETKAIKFIEQGIQRELTYKRADNSFSAFGDSDKAGSTWLTAFVVRSF 1070

Qy        997 LEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG-NKSPVTLTAYIVTS 1055

Db       1071 HHAKQYAFVDPNVISRAVAFLNSQQMESGAFAERGEVHHKDMQGGAQDGGVALTAFVLIS 1130

Qy       1056 LLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKEALNMLTWR- 1114

Db       1131 IL-----ENGMENGKAVTYLEKHLDEVSGNAYTMAVVAYALQLAKSKQAGKAFENLKKHK 1185

Qy       1115 -AEQEGGMQFWVSSE--SKLSDS----WQPRSLDIEVAAYALLSHFLQFQTSEGIPIMRW 1167

Db       1186 IVEKSGDVKFASAQKKVEKLKESRAYMFQARPVDIETTSYAVLSYLAQNQTSESLSIIRW 1245

Qy       1168 LSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTV-TGPSSPS------PLAV 1220

Db       1246 LVSQRNELGGFTSTQDTVMALQALSSYAAVTYSDKHTSQVTILNGKHTHSFDINIRNAIV 1305

Qy       1221 VQPM-------AVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVKEN 1273

Db       1306 LQSYQLSSLNDAVSINANGTGVVFAQLSYSY---------YRDSLNDDAPFFCSQEIKEI 1356

Qy       1274 KDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAISLSETVK-----KVEYDHG 1328

Db       1357 RAG-NRLQLDLCCNYTRPGKSNMALAEIDALSGYRFDAEQVHTLTSIEDLQRVEMEKDDT 1415

Qy       1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCD 1388

Db       1416 KMNVYFNPLGGRPVCLSLYSDVTYQVADQKPANFRLVDYYDPEEQLKMTYAAKQTRSLQE 1475

Qy       1389 LCSDVQGCRP 1398

Db       1476 KCG--EDCWP 1483


RESULT 4
```

```
US-10-292-081A-15
; Sequence 15, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
;  APPLICANT: Kenneth David Becker
;  APPLICANT: Gonul Velicelebi
;  APPLICANT:  Xin Wang
;  APPLICANT:  Randolph E. Tanzi
;  APPLICANT:  Lars Bertram
;  APPLICANT:  Aleister J. Saunders
;  TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLC
;  FILE REFERENCE: 37481-3323
;  CURRENT APPLICATION NUMBER: US/10/292,081A
;  CURRENT FILING DATE:  2002-11-08
;  PRIOR APPLICATION NUMBER: 60/337434
;  PRIOR FILING DATE: 2001-11-09
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 1474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-292-081A-15

  Query Match             19.6%;  Score 1441;  DB 12;  Length 1474;
  Best Local Similarity   29.2%;  Pred. No. 5.5e-109;
  Matches  453;  Conservative  264;  Mismatches  551;  Indels  284;  Gaps   60;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
Db         15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy         72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
Db         74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy        127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
Db        125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy        186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
Db        185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy        244 KHLNGTITAKYTYGKPVKGDVTLTFLPL---------------------------SFWGK 276
Db        244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303

Qy        277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
Db        304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy        332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
Db        348 TKLS-FVKVDSHFRQG---IPFFG---------------QVRLVDGKGVPIP----NKVI 384

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
Db        385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441

Qy        439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
Db        442 EEAHHTAY-----------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489

Qy        493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
Db        490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLM- 599
Db        550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608

Qy        600 -----NAS-----------------NDITMENVVHELELYNTGY-YLGMFMNS----FA 631
Db        609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDNEDCINRHNVYINGITYTPVSSTNEKDMYS 668

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
Db        669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725

Qy        681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
Db        726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782

Qy        738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
Db        783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841

Qy        798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
Db        842 EQAPHCICANGRQ-----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890

Qy        842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
Db        891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947

Qy        897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
Db        948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEIKSKAIGY 1007

Qy        957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
Db       1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067

Qy       1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
Db       1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127

Qy       1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
Db       1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186

Qy       1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
Db       1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245

Qy       1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
Db       1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
Db       1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353

Qy       1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
Db       1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413

Qy       1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
Db       1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465


RESULT 5
US-10-052-817-2
; Sequence 2, Application US/10052817
; Publication No. US20020114792A1
; GENERAL INFORMATION:
;  APPLICANT: Tanzi, Rudolph E.
;  APPLICANT:  Kovacs, Dora
;  APPLICANT:  Saunders, Aleister J.
;  TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
;  TITLE OF INVENTION:  Alzheimer's Disease
;  FILE REFERENCE: 0609.4460005
;  CURRENT APPLICATION NUMBER: US/10/052,817
;  CURRENT FILING DATE:  2002-01-23
```

```
; PRIOR APPLICATION NUMBER: 09/241,606
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: 09/148,503
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/093,297
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 1474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-052-817-2

  Query Match             19.6%;  Score 1441;  DB 14;  Length 1474;
  Best Local Similarity   29.1%;  Pred. No. 5.5e-109;
  Matches  451;  Conservative  266;  Mismatches  551;  Indels  284;  Gaps   59;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
              || :      :|: |:::|  | ::         |  || :    ||| | |      | ::
Db         15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy         72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
               : ||||  |     :|       ::| :|::|        ::| | ||:   |     | :  :
Db         74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy        127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
               :  ||:||||::||| | ||||:|::  :| |       : :: :|:|| | || | | :
Db        125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy        186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
              |:   :| |||  |   | : : ||        :| :   | | |:|||||||  |    ::
Db        185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy        244 KHLNGTITAKYTYGKPVKGDVTLTFLPL---------------------------SFWGK 276
               : :| ::     |||||| | ||::                              |: :
Db        244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303

Qy        277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
               |  || |::         : |||        |     :   | |  ||       :| ::
Db        304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy        332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
              | :|  |   :  |:|    | ||              :|   ||  : :      | ||
Db        348 TKLS-FVKVDSHFRQG---IPFFG--------------QVRLVDGKGVPIP----NKVI 384

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
               : :  |    |:  ||:    ||  ||  |    :   |  | ::  :          |:
Db        385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441

Qy        439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
               |    ||             |  ||||::: |:        :   |          :::|
Db        442 EEAHHTAY------------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489

Qy        493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
              ||:||  |::::| :|  |         ||          |:  ::   | | :::|  :  |
Db        490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ::| |   |  |:      ||: | :|  :: |:      ||::    | |:  : |||:|| ||
Db        550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608

Qy        601 ASNDITMENVVHEL-ELYNTGY--------------YLGMFMNS--------------FA 631
               :::  :| : | |  ||:                :::|                   ::
Db        609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 668

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
                :: ||  |:: : | :          |             |          :| : :
Db        669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725

Qy        681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
                ||   |||:||||||     :        |  |||||:|| | |   | :|| |||:::
Db        726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782

Qy        738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
               |    |:||||||: | :||||||| | |: |: |||        | | :| |  |   :
Db        783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841

Qy        798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
              |      | | | |           || : :  |  || :   ||:| :             |:
Db        842 EQAPHCICANGRQ-----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890

Qy        842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
                   | | : :|| : ||:|| : : ||  :     :      : ||   ||| |  | |
Db        891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947

Qy        897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
              :: :||:|| ::      :|::|||||||||: ||||||:|||| : :||  :| ||: :
Db        948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEIKSKAIGY 1007

Qy        957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
               :  ||||:| |:    |||:| ||   |    |:||:||||: | :|  || ||: :
Db       1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067

Qy       1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
               ||    || || |       | :::: ::|| :    |||:|||   :||         :  |: ::
Db       1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127

Qy       1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
              |||       |   |  | ||: ||  :  |:  | |||  |     |  ::    |
Db       1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186

Qy       1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
                :: :    ::|:  | ::|: :| ||: :|  |      |||  :     |::|:::|:|:
Db       1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245

Qy       1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------- 1222
              |||:||||| |||  ||| |||:  |   |||  ||| ||:  :
Db       1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
                  |    ::   |   |     | :: ||:                :  :| |   :  |:  :  |
Db       1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353

Qy       1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
               :    |     :  |  |  |      | ||:::| ::|||:        : :    |  | :  |
Db       1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413

Qy       1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
                : :||| |:      :      :::  | : : | | : ||||        |:   ||:
Db       1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465

RESULT 6
US-09-873-403-5
; Sequence 5, Application US/09873403
; Patent No. US20020028207A1
; GENERAL INFORMATION:
;  APPLICANT: Srivstava, Pramod K
;  TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC
;  TITLE OF INVENTION:  MOLECULES FOR IMMUNOTHERAPY
;  FILE REFERENCE: 8449-178
;  CURRENT APPLICATION NUMBER: US/09/873,403
;  CURRENT FILING DATE:  2001-06-04
;  PRIOR APPLICATION NUMBER: 09/625,139
;  PRIOR FILING DATE: 2000-07-25
;  PRIOR APPLICATION NUMBER: 60/209,266
;  PRIOR FILING DATE: 2000-06-02
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 1474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-873-403-5
```

```
Query Match             19.6%;  Score 1440;  DB 9;  Length 1474;
Best Local Similarity   29.1%;  Pred. No. 6.7e-109;
Matches  451;  Conservative  266;  Mismatches  551;  Indels  284;  Gaps   59;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
Db         15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy         72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
Db         74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy        127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
Db        125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy        186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
Db        185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy        244 KHLNGTITAKYTYGKPVKGDVTLTFLPL--------------------------SFWGK 276
Db        244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303

Qy        277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
Db        304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy        332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
Db        348 TKLS-FVKVDSHFRQG---IPFFG---------------QVRLVDGKGVPIP----NKVI 384

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
Db        385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441

Qy        439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
Db        442 EEAHHTAY------------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489

Qy        493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
Db        490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
Db        550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608

Qy        601 ASNDITMENVVHEL-ELYNTGY--------------YLGMFMNS--------------FA 631
Db        609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 668

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
Db        669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725

Qy        681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
Db        726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782

Qy        738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
Db        783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841

Qy        798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
Db        842 EQAPHCICANGRQ----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890

Qy        842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
Db        891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947

Qy        897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
Db        948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEVKSKAIGY 1007

Qy        957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
Db       1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067

Qy       1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
Db       1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127

Qy       1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
Db       1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186

Qy       1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
Db       1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245

Qy       1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
Db       1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
Db       1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353

Qy       1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
Db       1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413

Qy       1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
Db       1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465
```

RESULT 7
US-10-292-081A-10
; Sequence 10, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOI
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
;   LENGTH: 1474
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-292-081A-10

```
Query Match             19.6%;  Score 1440;  DB 12;  Length 1474;
Best Local Similarity   29.1%;  Pred. No. 6.7e-109;
Matches  451;  Conservative  266;  Mismatches  551;  Indels  284;  Gaps   59;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
Db         15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy         72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
Db         74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124
```

```
Qy        127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
Db        125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy        186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
Db        185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy        244 KHLNGTITAKYTYGKPVKGDVTLTFLPL---------------------------SFWGK 276
Db        244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303

Qy        277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
Db        304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy        332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
Db        348 TKLS-FVKVDSHFRQG---IPFFG--------------QVRLVDGKGVPIP----NKVI 384

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
Db        385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441

Qy        439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
Db        442 EEAHHTAY-----------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489

Qy        493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
Db        490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
Db        550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608

Qy        601 ASNDITMENVVHEL-ELYNTGY--------------YLGMFMNS--------------FA 631
Db        609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 668

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
Db        669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725

Qy        681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
Db        726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782

Qy        738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
Db        783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841

Qy        798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
Db        842 EQAPHCICANGRQ-----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890

Qy        842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
Db        891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947

Qy        897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
Db        948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEVKSKAIGY 1007

Qy        957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
Db       1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067

Qy       1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
Db       1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127

Qy       1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
Db       1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186

Qy       1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
Db       1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245

Qy       1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
Db       1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
Db       1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353

Qy       1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
Db       1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413

Qy       1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
Db       1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465
```

RESULT 8
US-10-292-081A-12
; Sequence 12, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLO
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-12

```
  Query Match             19.6%;  Score 1440;  DB 12;  Length 1474;
  Best Local Similarity   29.1%;  Pred. No. 6.7e-109;
  Matches  451;  Conservative  266;  Mismatches  551;  Indels  284;  Gaps   59;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
Db         15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy         72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
Db         74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy        127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
Db        125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy        186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
Db        185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy        244 KHLNGTITAKYTYGKPVKGDVTLTFLPL---------------------------SFWGK 276
Db        244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303
```

```
Qy        277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
              |  || |::           : |||          |      |  |  ||      :| ::
Db        304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy        332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
              | :|  |  :  |:|     | ||                 :|   ||  :  :     | ||
Db        348 TKLS-FVKVDSHFRQG---IPFFG---------------QVRLVDGKGVPIP----NKVI 384

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP---------ILEDS 438
               : : |    |:  ||:   :    ||  || |  :  |  | :: :          : |:
Db        385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441

Qy        439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
               |     ||              |   ||||:::: |:      :   |          :::|
Db        442 EEAHHTAY------------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489

Qy        493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
              ||:||  |::::|  :|  |          ||          .|:  ::   || | :::| :  |
Db        490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ::|  |   |       ||:  | :|   :: |:      ||::     |  :   |||:|| ||
Db        550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608

Qy        601 ASNDITMENVVHEL-ELYNTGY--------------YLGMFMNS--------------FA 631
               :::  :| :  |  |   ||:                 :::|                 ::
Db        609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 668

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
               :: ||   |:: :  |   :          |            |     |       :| : :
Db        669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725

Qy        681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
                  ||    |||:|||||||        :         |   |||||:|| |   |  :||| |||:::
Db        726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782

Qy        738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
              |  |:|||||:  |  :|||||||:  |  |  |   |:  :  |  |||   |   :
Db        783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841

Qy        798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
              |     | | | |             || : :  |  ||  :  ||:|  :               |:
Db        842 EQAPHCICANGRQ-----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890

Qy        842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
                       | |  :  :||:  ||:||   :  : ||   :     :     ||    |||  | |
Db        891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947

Qy        897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
              :: :||:|| ::       :|::||||||||||: ||||||:||||  : :|||   :| ||: :
Db        948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEVKSKAIGY 1007

Qy        957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
              :    |||:| |:        |      |       | :|||:|||:|  |  :|  ||  :|   :
Db       1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067

Qy       1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
               ||    || || |      | ::::  :||  |  |||:|||  :||        :   |: ::
Db       1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127

Qy       1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
              |||        |   |  |  ||  ||: ||  :   |:  |  ||   |   |   ::    |
Db       1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186

Qy       1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
                :: :     ::|:  |  ::|:  :| ||: :|  |    ||| :      |::|:::|:|:
Db       1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245

Qy       1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
              |||:||||| |||  |||::  |    | |    ||   ||   :   |
Db       1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
                  |   ::   |  |   | :: ||:                :  :|  |   :  |:   :  |
Db       1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353

Qy       1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
                :   |    |  |  |     | ||:::| ::|||:      :  :   |  |  :  |
Db       1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413

Qy       1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
               : :||| |:     :     :::    | : :  |  | : ||||   |:  ||:
Db       1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465
```

RESULT 9
US-10-292-081A-13
; Sequence 13, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOB
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-13

```
  Query Match             19.6%;  Score 1440;  DB 12;  Length 1474;
  Best Local Similarity   29.1%;  Pred. No. 6.7e-109;
  Matches  451;  Conservative  266;  Mismatches  551;  Indels  284;  Gaps   59;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
               |.| :     :|:  |:::|  | ::             |  ||  :       ||| |  |       | ::
Db         15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy         72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
               : ||||    |      :|         ::| :|::|          ::| |  | ||:    |     |  :   :
Db         74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy        127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
                :    |||:|:||| |  ||||:|::  :|      |  :  |:|||  |  |  |   |  |  |  |
Db        125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy        186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
                |:   :|  |||  |     |  :  : ||          :|  :      |  |  | |:||||||  |       ::
Db        185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy        244 KHLNGTITAKYTYGKPVKGDVTLTFLPL--------------------------SFWGK 276
               : :|  ::     ||||||  | ||::                                 |:  :
Db        244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303

Qy        277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
              |  || |::           : |||          |      |  |  ||      :| ::
Db        304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy        332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
              | :|  |  :  |:|     | ||                 :|   ||  :  :     | ||
Db        348 TKLS-FVKVDSHFRQG---IPFFG---------------QVRLVDGKGVPIP----NKVI 384

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
               : : |    |:  ||:   :    ||  || |  :  |  | :: :          : |:
Db        385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441
```

```
Qy     439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
Db     442 EEAHHTAY------------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489

Qy     493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
Db     490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549

Qy     541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
Db     550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608

Qy     601 ASNDITMENVVHEL-ELYNTGY--------------YLGMFMNS--------------FA 631
Db     609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 668

Qy     632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
Db     669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725

Qy     681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
Db     726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782

Qy     738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
Db     783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841

Qy     798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
Db     842 EQAPHCICANGRQ-----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890

Qy     842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
Db     891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947

Qy     897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
Db     948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEVKSKAIGY 1007

Qy     957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
Db    1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067

Qy    1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
Db    1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127

Qy    1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
Db    1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186

Qy    1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
Db    1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245

Qy    1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
Db    1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304

Qy    1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
Db    1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353

Qy    1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
Db    1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413

Qy    1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
Db    1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465
```

```
RESULT 10
US-10-331-496A-38
; Sequence 38, Application US/10331496A
; Publication No. US20030228305A1
; GENERAL INFORMATION:
;  APPLICANT: FRANTZ,GRETCHEN
;  APPLICANT:  HILLAN,KENNETH J.
;  APPLICANT:  PHILLIPS,HEIDI S.
;  APPLICANT:  POLAKIS,PAUL
;  APPLICANT:  SMITH,VICTORIA
;  APPLICANT:  SPENCER,SUSAN D.
;  APPLICANT:  WILLIAMS,P. MICKEY
;  APPLICANT:  WU,THOMAS D.
;  APPLICANT:  ZHANG,ZEMIN
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF TUMOR
;  FILE REFERENCE: P5014R1-PCT
;  CURRENT APPLICATION NUMBER: US/10/331,496A
;  CURRENT FILING DATE:  2002-12-30
;  PRIOR APPLICATION NUMBER: US 60/345,444
;  PRIOR FILING DATE: 2002-01-02
;  PRIOR APPLICATION NUMBER: US 60/351,885
;  PRIOR FILING DATE: 2002-01-25
;  PRIOR APPLICATION NUMBER: US 60/360,066
;  PRIOR FILING DATE: 2002-02-25
;  PRIOR APPLICATION NUMBER: US 60/362,004
;  PRIOR FILING DATE: 2002-03-05
;  PRIOR APPLICATION NUMBER: US 60/366,869
;  PRIOR FILING DATE: 2002-03-20
;  PRIOR APPLICATION NUMBER: US 60/366,284
;  PRIOR FILING DATE: 2002-03-21
;  PRIOR APPLICATION NUMBER: US 60/368,679
;  PRIOR FILING DATE: 2002-03-28
;  PRIOR APPLICATION NUMBER: US 60/404,809
;  PRIOR FILING DATE: 2002-08-19
;  PRIOR APPLICATION NUMBER: US 60/405,645
;  PRIOR FILING DATE: 2002-08-21
;  NUMBER OF SEQ ID NOS: 95
; SEQ ID NO 38
;   LENGTH: 1474
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-331-496A-38

  Query Match             19.6%;  Score 1440;  DB 12;  Length 1474;
  Best Local Similarity   29.1%;  Pred. No. 6.7e-109;
  Matches  451;  Conservative  266;  Mismatches  551;  Indels  284;  Gaps   59;

Qy      12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
Db      15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy      72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
Db      74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy     127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
Db     125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy     186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
Db     185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy     244 KHLNGTITAKYTYGKPVKGDVTLTFLPL---------------------------SFWGK 276
Db     244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303

Qy     277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
Db     304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy     332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
```

```
               | :|   |   :   |:|   .|.||                |   ||  : :     | ||
Db        348 TKLS-FVKVDSHFRQG---IPPFG--------------QVRLVDGKGVPIP----NKVI 384

Qy        392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP---------ILEDS 438
               : : |     |:  ||:    :     ||    || |   :  | | :: :            : |:
Db        385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441

Qy        439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
               |     ||              |||:::  |:        :   |            :::|
Db        442 EEAHHTAY-----------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489

Qy        493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
              ||:||  |:::::| :|  |       ||          |:  ::    | | :::| :   |
Db        490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
              ::|  |   |   |:     ||: | :|  :: |:     ||::    |  |:  : |||:|| ||
Db        550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608

Qy        601 ASNDITMENVVHEL-ELYNTGY-------------YLGMFMNS-------------FA 631
                 :::  :| : |  |    ||:                  :::|               ::
Db        609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 668

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAE----------YAERFMEENEGHIVDIHDFSL 680
               :: ||    |:: : |  :     |                  |     |      :| : :
Db        669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725

Qy        681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
                  ||    |||:|||||||         :          |||||:|| |  |    :||| |||:::
Db        726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782

Qy        738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
               |   |:|||||: | :||||||| |  |: |: |||        | | :| |   |   :
Db        783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841

Qy        798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS----------PTA 841
               |      | | | |               || : :  |   || :   ||:|  :                |:
Db        842 EQAPHCICANGRQ----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890

Qy        842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
                   |  | : :||: ||:|  |   : ||   :         :  ||  |||  |   |
Db        891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947

Qy        897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
              :: :||:||  ::       :|::|||||||||:  |||||:||||  : :|||   :|  ||: :
Db        948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEVKSKAIGY 1007

Qy        957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
              :  ||||:|  |:   |||:| ||   |    :  |:|||:||||:  | :|   || ||:   :  :
Db       1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067

Qy       1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
               ||   || || |    | :::: ::|| :  |||:|||  :||        :   |: ::
Db       1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127

Qy       1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
              |||        |    |   |  |:| ||: ||   :  |:  | |  |   |   |  ::     |
Db       1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186

Qy       1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
               :: :   : ::|: :  |::|: :| ||: :|  :     |||     |::|:::|:|:
Db       1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245

Qy       1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------- 1222
              |||:||||| ||| |||::    |   ||       || ||  :   :   |
Db       1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
                   |    ::   | |   | :: ||:             :  :|  |   : |:   :  |
Db       1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353

Qy       1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
                :   |   :  | | |    | ||:::| ::|||:    : :    |  | : |
Db       1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413

Qy       1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNS 1380
               : :||| |:     :     :::  | : : | | : ||||    |:  ||:
Db       1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEYNA 1465
```

RESULT 11
US-09-756-247-4
; Sequence 4, Application US/09756247
; Publication No. US20030180722A1
; GENERAL INFORMATION:
; APPLICANT: Godbole, Shubhada D
; APPLICANT: Boyle, Bryan J
; APPLICANT: Mize, Nancy K
; APPLICANT: Deng, Cenhua
; APPLICANT: Goodrich, Ryle
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Zhou, Ping
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Yeung, George
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE POI
; TITLE OF INVENTION: AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-31CIP
; CURRENT APPLICATION NUMBER: US/09/756,247
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/684,711
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1508
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-756-247-4

```
  Query Match             19.6%;  Score 1437.5;  DB 12;  Length 1508;
  Best Local Similarity   27.1%;  Pred. No. 1.1e-108;
  Matches  422;  Conservative  274;  Mismatches  552;  Indels  307;  Gaps   45;

Qy         20 LAVAPG-----PRFLVTAPGIIR-PGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSV 73
              ||::|       | :|||  |   :  |    ::       |:   |     ||   :     :| |  :
Db         10 LALSPAIAEELPNYLVTLPARLNFP----SVQKVCLDLSPGYSDVKFTVTLETKDKTQKL 65

Qy         74 LEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRISVFI 133
              || |: ::         |    |        :|:  :||:|     :  |  |         ::  :  :   |:
Db         66 LEYSGLKKRHLHCISFLVPPPAGGTEEVATIRVSG-VGNNISFEEKKKVLIQRQGNGTFV 124

Qy        134 QTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQSQDLGVISKT 192
              |||| ||  |  |:| |||||: |:|  |           ::: ::|| || | |||    :  |::  :
Db        125 QTDKPLYTPGQQVYFRIVTMDSNFVPVNDKYSMVELQDPNSNRIAQWLEVVPEQGIVDLS 184

Qy        193 FQLSSHPILGDWSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITA 252
              |||:    :||  ::    | |   :       :  :| |||||||:|: :    |     :       |
Db        185 FQLAPEAMLGTYT--VAVAEGKTFGTFSVEEYVLPKFKVEVVEPKELSTVQESFLVKICC 242

Qy        253 KYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEY 312
              :||||||: | | ::               :| |        .         : : : | : : |     ||
Db        243 RYTYGKPMLGAVQVSVC------QKAN-----------TYWYREVEREQLPDKCRNLSGQ 285

Qy        313 LD----LSSP-------------GPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFF 354
                |      |:|                 : |: ||  |   ||:  | : |::            :  |
Db        286 TDKTGCFSAPVDMATFDLIGYAYSHQINIVATVVEEGTGVEANATQNIYISPQMGSMTFE 345
```

```
Qy        355 DYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI-------------TVTQRN---- 397
Db        346 DTSNFYHPNFPFSGKIRVRGHDDSFL-----KNHLVFLVIYGTNGTFNQTLVTDNNGLAP 400

Qy        398 ---YTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIE--------FPILEDSSELQLKAY 446
Db        401 FTLETSGWNGTD---------------VSLEGKFQMEDLVYNPEQVPRYYQNAYLHLRPF 445

Qy        447 FLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVV---------SGNKRLKELS 497
Db        446 YSTTRSFLGIHRL--------------NGPLKCGQPQEVLVDYYIDPADASPDQEI-SFS 490

Qy        498 YMVVSRGQLVAVGKQN----------STMFSLTPENSWTPKACVIVYYIEDDGEIISDVL 547
Db        491 YYLIGKGSLVMEGQKHLNSKKKGLKASFSLSLTFTSRLAPDPSLVIYAIFPSGGVVADKI 550

Qy        548 KIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITM 607
Db        551 QFSVGMCFDNQVSLGFSPSQQLPGAEVELQLQAA-PGSLCALRAVDESVLLLRP------ 603

Qy        608 ENVVHELELYNTGYYLGMF----------------------------------------- 626
Db        604 -----DRELSNRSVY-GMFPFWYGHYPYQVAEYDQCPVSGPWDFPQPLIDPMPQGHSSQR 657

Qy        627 -----------MNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYA---------ERFME 666
Db        658 SIIWRPSFSEGTDLFSFFRDVGLKILSNAKIKKP-VDCSHRSPEYSTAMGGGGHPEAFES 716

Qy        667 ENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFV 726
Db        717 STPLHQAE--------DSQVRQYFPETWLWDLFPIGNSGKEAVHVTVPDAITEWKAMSFC 768

Qy        727 ISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEK 786
Db        769 TSQSRGFGLSPT-VGLTAFKPFFVDLTLPYSVVRGESFRLTATIFNYLKDCIRVQTDLAK 827

Qy        787 SDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTAL--------- 837
Db        828 SHEYQLESWADS------QTSSCLCADDAKTHHWNITAVKLGHINFTISTKILDSNEPCG 881

Qy        838 -------SPTASDAVTQMILVKAEG--IEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNT 888
Db        882 GQKGFVPQKGRSDTLIKPVLVKPEGVLVEKTHS-----SLLCPKGKVASESVSLELPVDI 936

Qy        889 VTGSERVQITAIG---------------------DVLGPSINGLASLIRMPYGCGEQNMI 927
Db        937 VPDSTKAYVTVLGKQLEILDSERKRRMEAAKVWRDIMGTALQNLDGLVQMPSGCGEQNMV 996

Qy        928 NFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTW 987
Db        997 LFAPIIYVLQYLEKAGLLTEEIRSRAVGFLEIGYQKELMYKHSNGSYSAFGERDGNGNTW 1056

Qy        988 LSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVT 1047
Db       1057 LTAFVTKCFGQAQKFIFIDPKNIQDALKWMAGNQLPSGCYANVGNLLHTAMKGGVDDEVS 1116

Qy       1048 LTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKE- 1106
Db       1117 LTAYVTAALLEMGKDVDDPMVSQGLRCLKNS-ATSTTNLYTQALLAYIFSLAGEMDIRNI 1175

Qy       1107 ALNMLTWRAEQEGGMQFWVSSESKLSDS--W-QPRSLDIEVAAYALLSH-----FLQFQT 1158
Db       1176 LLKQLDQQAIISGESIYWSQKPTPSSNASPWSEPAAVDVELTAYALLAQLTKPSLTQKEI 1235

Qy       1159 SEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFA--ALMNTERTNIQVTVT------ 1210
Db       1236 AKATSIVAWLAKQHNAYGGFSSTQDTVVALQALAKYATTAYMPSEEINLVVKSTENFQRT 1295

Qy       1211 -GPSSPSPLAVVQ------PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEA 1263
Db       1296 FNIQSVNRLVFQQDTLPNVPGMYTLEASGQGCVYVQTVLRYNILPP---------TNMKT 1346

Qy       1264 FDLDVAVK----ENKDDLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGF--MVPSEAISL 1316
Db       1347 FSLSVEIGKARCEQPTSPRSLTLTIHTSYVGSRSSSNMAIVEVKMLSGFSPMEGTNQLLL 1406

Qy       1317 SE-TVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEP 1370
Db       1407 QQPLVKKVEFGTDTLNIYLDELIKNTQTYTFTISQSVLVTNLKPATIKVYDYYLP 1461
```

```
RESULT 12
US-09-756-247-23
; Sequence 23, Application US/09756247
; Publication No. US20030180722A1
; GENERAL INFORMATION:
;  APPLICANT:  Godbole, Shubhada D
;  APPLICANT:  Boyle, Bryan J
;  APPLICANT:  Mize, Nancy K
;  APPLICANT:  Deng, Cenhua
;  APPLICANT:  Goodrich, Ryle
;  APPLICANT:  Arterburn, Matthew C
;  APPLICANT:  Zhou, Ping
;  APPLICANT:  Tang, Y. Tom
;  APPLICANT:  Liu, Chenghua
;  APPLICANT:  Yeung, George
;  APPLICANT:  Drmanac, Radoje T
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE PO
;  TITLE OF INVENTION:  AND POLYNUCLEOTIDES
;  FILE REFERENCE:  HYS-31CIP
;  CURRENT APPLICATION NUMBER: US/09/756,247
;  CURRENT FILING DATE:  2001-01-08
;  PRIOR APPLICATION NUMBER:   09/649,167
;  PRIOR FILING DATE:   2000-08-23
;  PRIOR APPLICATION NUMBER:   09/540,217
;  PRIOR FILING DATE:   2000-03-31
;  PRIOR APPLICATION NUMBER:   09/684,711
;  PRIOR FILING DATE:   2000-10-06
;  PRIOR APPLICATION NUMBER:   09/560,875
;  PRIOR FILING DATE:   2000-04-27
;  PRIOR APPLICATION NUMBER:   09/496,914
;  PRIOR FILING DATE:   2000-02-03
;  NUMBER OF SEQ ID NOS:  41
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 23
;   LENGTH: 1450
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-756-247-23

  Query Match             19.5%;  Score 1433;  DB 12;  Length 1450;
  Best Local Similarity   29.0%;  Pred. No. 2.5e-108;
  Matches  450;  Conservative  265;  Mismatches  551;  Indels  284;  Gaps   59;

Qy         12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
Db          2 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 60

Qy         72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
Db         61 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 111

Qy        127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
Db        112 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 171

Qy        186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
Db        172 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 230

Qy        244 KHLNGTITAKYTYGKPVKGDVTLTFLPL-------------------------SFWGK 276
Db        231 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 290

Qy        277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
```

```
Db    291 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 334

Qy    332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
Db    335 TKLS-FVKVDSHFRQG---IPFFG--------------QVRLVDGKGVPIP----NKVI 371

Qy    392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
Db    372 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 428

Qy    439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
Db    429 EEAHHTAY-----------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 476

Qy    493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
Db    477 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 536

Qy    541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMN 600
Db    537 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 595

Qy    601 ASNDITMENVVHEL-ELYNTGY--------------YLGMFMNS--------------FA 631
Db    596 PDAELSASSVYNLLPEKDLTGFPGPLNDQDDEDCINRHNVYINGITYTPVSSTNEKDMYS 655

Qy    632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
Db    656 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 712

Qy    681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
Db    713 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 769

Qy    738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
Db    770 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 828

Qy    798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
Db    829 EQAPHCICANGRQ-----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 877

Qy    842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
Db    878 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 934

Qy    897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
Db    935 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEVKSKAIGY 994

Qy    957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
Db    995 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1054

Qy   1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
Db   1055 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1114

Qy   1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
Db   1115 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1173

Qy   1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
Db   1174 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1232

Qy   1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
Db   1233 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1291

Qy   1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
Db   1292 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1340

Qy   1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
Db   1341 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1400

Qy   1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSY 1378
Db   1401 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYETDEFAIAEY 1450
```

RESULT 13
US-09-756-247-25
; Sequence 25, Application US/09756247
; Publication No. US20030180722A1
; GENERAL INFORMATION:
; APPLICANT: Godbole, Shubhada D
; APPLICANT: Boyle, Bryan J
; APPLICANT: Mize, Nancy K
; APPLICANT: Deng, Cenhua
; APPLICANT: Goodrich, Ryle
; APPLICANT: Arterburn, Matthew C
; APPLICANT: Zhou, Ping
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Yeung, George
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE POI
; TITLE OF INVENTION: AND POLYNUCLEOTIDES
; FILE REFERENCE: HYS-31CIP
; CURRENT APPLICATION NUMBER: US/09/756,247
; CURRENT FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/684,711
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
;   LENGTH: 1491
;   TYPE: PRT
;   ORGANISM: hono sapiens
US-09-756-247-25

```
  Query Match             19.5%;  Score 1433;  DB 12;  Length 1491;
  Best Local Similarity   27.1%;  Pred. No. 2.6e-108;
  Matches  419;  Conservative  272;  Mismatches  551;  Indels  302;  Gaps   44;

Qy     26 PRFLVTAPGIIR-PGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGS 84
Db      4 PNYLVTLPARLNFP----SVQKVCLDLSPGYSDVKFTVTLETKDKTQKLLEYSGLKKRHL 59

Qy     85 FKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQ 144
Db     60 HCISFLVPPPAGGTEEVATIRVSG-VGNNISFEEKKKVLIQRQGNGTFVQTDKPLYTPGQ 118

Qy    145 EVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGD 203
Db    119 QVYFRIVTMDSNFVPVNDKYSMVELQDPNSNRIAQWLEVVPEQGIVDLSFQLAPEAMLGT 178

Qy    204 WSIQVQVNDQTYYQSFQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGD 263
Db    179 YT--VAVAEGKTFGTFSVEEYVLPKFKVEVVEPKELSTVQESFLVKICCRYTYGKPMLGA 236

Qy    264 VTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLD----LSSP- 318
Db    237 VQVSVC------QKAN-----------TYWYREVEREQLPDKCRNLSGQTDKTGCFSAPV 279
```

```
Qy    319 ------------GPVEILTTVTESVTGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLN 365
Db    280 DMATFDLIGYAYSHQINIVATVVEEGTGVEANATQNIYISPQMGSMTFEDTSNFYHPNFP 339

Qy    366 FTATVKVTRADGNQLTLEERRNNVVI-------------TVTQRN-------YTEYWSGS 405
Db    340 FSGKIRVRGHDDSFL-----KNHLVFLVIYGTNGTFNQTLVTDNNGLAPFTLETSGWNGT 394

Qy    406 NSGNQKMEAVQKINYTVPQSGTFKIE--------FPILEDSSELQLKAYFLGSKSSMAVH 457
Db    395 D---------------VSLEGKFQMEDLVYNPEQVPRYYQNAYLHLRPFYSTTRSFLGIH 439

Qy    458 SLFKSPSKTYIQLKTRDENIKVGSPFELVV---------SGNKRLKELSYMVVSRGQLVA 508
Db    440 RL-------------NGPLKCGQPQEVLVDYYIDPADASPDQEI-SFSYYLIGKGSLVM 484

Qy    509 VGKQN----------STMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNK 558
Db    485 EGQKHLNSKKKGLKASFSLSLTFTSRLAPDPSLVIYAIFPSGGVVADKIQFSVGMCFDNQ 544

Qy    559 IKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYN 618
Db    545 VSLGFSPSQQLPGAEVELQLQAA-PGSLCALRAVDESVLLLRP-----------DRELSN 592

Qy    619 TGYYLGMF---------------------------------------------------- 626
Db    593 RSVY-GMFPFWYGHYPYQVAEYDQCPVSGPWDFPQPLIDPMPQGHSSQRSIIWRPSFSEG 651

Qy    627 MNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYA---------ERFMEENEGHIVDIHD 677
Db    652 TDLFSFFRDVGLKILSNAKIKKP-VDCSHRSPEYSTAMGGGGHPEAFESSTPLHQAE--- 707

Qy    678 FSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
Db    708 -----DSQVRQYFPETWLWDLFPIGNSGKEAVHVTVPDAITEWKAMSFCTSQSRGFGLSP 762

Qy    738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
Db    763 T-VGLTAFKPFFVDLTLPYSVVRGESFRLTATIFNYLKDCIRVQTDLAKSHEYQLESWAD 821

Qy    798 EINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTAL----------------SPTA 841
Db    822 S------QTSSCLCADDAKTHHWNITAVKLGHINFTISTKILDSNEPCGGQKGFVPQKGR 875

Qy    842 SDAVTQMILVKAEG--IEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITA 899
Db    876 SDTLIKPVLVKPEGVLVEKTHS-----SLLCPKGKVASESVSLELPVDIVPDSTKAYVTV 930

Qy    900 IG--------------------DVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDY 938
Db    931 LGKQLEILDSERKRRMEAAKVWRDIMGTALQNLDGLVQMPSGCGEQNMVLFAPIIYVLQY 990

Qy    939 LTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLE 998
Db    991 LEKAGLLTEEIRSRAVGFLEIGYQKELMYKHSNGSYSAFGERDGNGNTWLTAFVTKCFGQ 1050

Qy    999 ADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLG 1058
Db   1051 AQKFIFIDPKNIQDALKWMAGNQLPSGCYANVGNLLHTAMKGGVDDEVSLTAYVTAALLE 1110

Qy   1059 YRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKE-ALNMLTWRAEQ 1117
Db   1111 MGKDVDDPMVSQGLRCLKNS-ATSTTNLYTQALLAYIFSLAGEMDIRNILLKQLDQQAII 1169

Qy   1118 EGGMQFWVSSESKLSDS--W-QPRSLDIEVAAYALLSH-----FLQFQTSEGIPIMRWLS 1169
Db   1170 SGESIYWSQKPTPSSNASPWSEPAAVDVELTAYALLAQLTKPSLTQKEIAKATSIVAWLA 1229

Qy   1170 RQRNSLGGFASTQDTTVALKALSEFA--ALMNTERTNIQVTVT-------GPSSPSPLAV 1220
Db   1230 KQHNAYGGFSSTQDTVVALQALAKYATTAYMPSEEINLVVKSTENFQRTFNIQSVNRLVF 1289

Qy   1221 VQ------PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK--- 1271
Db   1290 QQDTLPNVPGMYTLEASGQGCVYVQTVLRYNILPP---------TNMKTFSLSVEIGKAR 1340

Qy   1272 -ENKDDLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGF--MVPSEAISLSE-TVKKVEYD 1326
Db   1341 CEQPTSPRSLTLTIHTSYVGSRSSSNMAIVEVKMLSGFSPMEGTNQLLLQQPLVKKVEFG 1400

Qy   1327 HGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEP 1370
Db   1401 TDTLNIYLDELIKNTQTYTFTISQSVLVTNLKPATIKVYDYYLP 1444
```

RESULT 14
US-10-292-081A-9
; Sequence 9, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLO
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1500
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-9

Query Match 19.4%; Score 1428; DB 12; Length 1500;
Best Local Similarity 28.7%; Pred. No. 6.7e-108;
Matches 458; Conservative 270; Mismatches 567; Indels 300; Gaps 62;

```
Qy     12 LLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTIGVELLEHCPSQVTVKAELLKTASNLTV 71
Db     15 LLVLLPTDASVSGKPQYMVLVPSLLHT-ETTEKGCVLLSYLNETVTVSASLESVRGNRSL 73

Qy     72 -SVLEAEG-VFEKGSFKTLTLPSLPLNSADE---IYELRVTGRTQDEILFSNSTRLSFET 126
Db     74 FTDLEAENDVLHCVAF------AVPKSSSNEEVMFLTVQVKGPTQE---FKKRTTVMVKN 124

Qy    127 KRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSD 185
Db    125 EDSLVFVQTDKSIYKPGQTVKFRVVSMDENFHPLNELIPLVYIQDPKGNRIAQWQSFQLE 184

Qy    186 LGVISKTFQLSSHPILGDWSIQVQVND--QTYYQSFQVSEYVLPKFEVTLQTPLYCSMNS 243
Db    185 GGLKQFSFPLSSEPFQGSYKVVVQKKSGGRTEH-PFTVEEFVLPKFEVQVTVPKIITILE 243

Qy    244 KHLNGTITAKYTYGKPVKGDVTLTFLPL---------------------------SFWGK 276
Db    244 EEMNVSVCGLYTYGKPVPGHVTVSICRKYSDASDCHGEDSQAFCEKFSGQLNSHGCFYQQ 303

Qy    277 KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILT-----TVTESV 331
Db    304 VK--TKVFQLKR------KEYEMK--------LHTEAQIQEEGTVVELTGRQSSEITRTI 347

Qy    332 TGISRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI 391
Db    348 TKLS-FVKVDSHFRQG---IPFFG---------------QVRLVDGKGVPIP----NKVI 384

Qy    392 TVTQRNYTEYWSGSNSGNQKMEAVQ-KINYT--VPQSGTFKIEFP----------ILEDS 438
Db    385 FI-RGNEANYY--SNATTDEHGLVQFSINTTNVMGTSLTVRVNYKDRSPCYGYQWVSEEH 441
```

```
Qy        439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGS----PFELVVSGNK--R 492
              |     ||           |  |||:::  |:       :   |        :::|
Db        442 EEAHHTAY-----------LVFSPSKSFVHLEPMSHELPCGHTQTVQAHYILNGGTLLG 489

Qy        493 LKELS--YMVVSRGQLVAVG------KQNST----MFSLTPENSWTPKACVIVYYIEDDG 540
              ||:||  |:::::| :|   |       ||          |:  ::   | | :::| :   |
Db        490 LKKLSFYYLIMAKGGIVRTGTHGLLVKQEDMKGHFSISIPVKSDIAPVARLLIYAVLPTG 549

Qy        541 EIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLM- 599
              ::| |   |   |:    ||: | :|  :: |:    ||::   | |:  : |||:|| ||
Db        550 DVIGDSAKYDVENCLANKVDLSFSPSQSLPASHAHLRVTAA-PQSVCALRAVDQSVLLMK 608

Qy        600 -----NAS-----------------NDITMENVVHELELYNTGY-YLGMFMNS----FA 631
                   :||                 ||  |: ::  :|  |   |   :   :    ::
Db        609 PDAELSASSVYNLLPEKDLTGFPGPLNDQDNEDCINRHNVYINGITYTPVSSTNEKDMYS 668

Qy        632 VFQECGLWVLTDANLTKDYIDGVYDNAE-----------YAERFMEENEGHIVDIHDFSL 680
                :: ||   |:: : |  :          |             |     |       :| : :
Db        669 FLEDMGLKAFTNSKIRKPKMCPQLQQYEMHGPEGLRVGFYESDVMGRGHARLVHVEE--- 725

Qy        681 GSSPH---VRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTT 737
                 ||    |||:||||||      :        |   |||||:|| |  |   | :||| |||:::
Db        726 ---PHTETVRKYFPETWIWDLVVVNSAGVAEVGVTVPDTITEWKAGAFCLSEDAGLGISS 782

Qy        738 TPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSS 797
              |    |:||||||: | :|||||||| | |: |: |||       | | :| |  |  |  :
Db        783 T-ASLRAFQPFFVELTMPYSVIRGEAFTLKATVLNYLPKCIRVSVQLEASPAFLAVPVEK 841

Qy        798 E-----INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALS-----------PTA 841
              |      | | | |                 || : :  |   || :    ||:| :               |:
Db        842 EQAPHCICANGRQ----------TVSWAVTPKSLGNVNFTVSAEALESQELCGTEVPSV 890

Qy        842 -----SDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQ 896
                   | |  : :||: ||:|| : : ||  :    :    :  :||       ||| |  |  |
Db        891 PEHGRKDTVIKPLLVEPEGLEKETTFNSLLCPSGGEVS---EELSLKLPPNVVEESARAS 947

Qy        897 ITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSF 956
              :: :||:|| ::    :|::||||||||||: ||||||:||||  : :|||   :| ||: :
Db        948 VSVLGDILGSAMQNTQNLLQMPYGCGEQNMVLFAPNIYVLDYLNETQQLTPEIKSKAIGY 1007

Qy        957 MRQGYQRELLYQREDGSFSAFG-NYDPS-GSTWLSAFVLRCFLEADPYIDIDQNVLHRTY 1014
              :  ||||:| |:   |||:| ||   |   : |:|||:||||: | :|  || ||:   : :
Db       1008 LNTGYQRQLNYKHYDGSYSTFGERYGRNQGNTWLTAFVLKTFAQARAYIFIDEAHITQAL 1067

Qy       1015 TWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHF 1074
               ||  || || ||   |  ::::  ::|| :   |||:|||  :|      :   |: ::
Db       1068 IWLSQRQKDNGCFRSSGSLLNNAIKGGVEDEVTLSAYITIALLEIPLTVTHPVVRNALFC 1127

Qy       1075 LES------EFSRGISDNYTLALITYALSSVGS-PKAKEALNMLTWRAEQEGGMQFW--- 1124
              |||      |  |   || ||: ||  :  |:  | |||  |   |   ::        |
Db       1128 LESAWKTAQEGDHG-SHVYTKALLAYAFALAGNQDKRKEVLKSLNEEAVKKDNSVHWERP 1186

Qy       1125 VSSESKLSDSWQPR--SLDIEVAAYALLSHFLQFQ---TSEGIP----IMRWLSRQRNSL 1175
                :: :   ::|:  | ::|: :| ||: :|  |   |||  :      |::|:::|:|:
Db       1187 QKPKAPVGHFYEPQAPSAEVEMTSYVLLA-YLTAQPAPTSEDLTSATNIVKWITKQQNAQ 1245

Qy       1176 GGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSPSPLAVVQ------------ 1222
              |||:||||| ||| |||::  |   |  ||       ||  ||  :  :  |
Db       1246 GGFSSTQDTVVALHALSKYGAATFT-RTGKAAQVTIQSSGTFSSKFQVDNNNRLLLQQVS 1304

Qy       1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK---ENKD 1275
                  |   ::   | |    | :: ||:              :  :| |    :  |:   :  |
Db       1305 LPELPGEYSMKVTGEGCVYLQTSLKYNI-----------LPEKEEFPFALGVQTLPQTCD 1353

Qy       1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLLSGFMVPSEAISL---SETVKKVEYDHG 1328
              :   |    :   |  |  |    | ||:::| ::|||:      : :    |  | : |
Db       1354 EPKAHTSFQISLSVSYTGSRSASNMAIVDVKMVSGFIPLKPTVKMLERSNHVSRTEVSSN 1413

Qy       1329 KLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRRQAVRSYNSEVKLSSCD 1388
               : :||| |:     :    :::  | :  :  | |  |  : ||||             :::| |
Db       1414 HVLIYLDKVSNQTLSLFFTVLQDVPVRDLKPAIVKVYDYYE---------TGDLQLLSTM 1464

Qy       1389 LCS-------DVQGCRPCEDGASGSHHHSSVIFIF 1416
                 | :        :        |  |   |  |   :|::
Db       1465 LLAAKILEMLEDHKAEKCFAGVLFSELHRRHVFLY 1499
```

RESULT 15
US-09-981-151A-10
; Sequence 10, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shimkets, Richard A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Esha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1492
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-981-151A-10

```
  Query Match             19.3%;  Score 1421.5;  DB 12;  Length 1492;
  Best Local Similarity   27.2%;  Pred. No. 2.3e-107;
  Matches  419;  Conservative  269;  Mismatches  559;  Indels  293;  Gaps   44;

Qy         20 LAVAPG-----PRFLVTAPGIIR-PGGNVTIGVELLEHCPSQVTVKAELLKTASNLTVSV 73
              ||::|      | :||| |  :  |      ::     |:  |    ||  :      :  |  :
Db         10 LALSPAIAEELPNYLVTLPARLNFP----SVQKVCLDLSPGYSDVKFTVTLETKDKTQKL 65
```

```
Qy     74 LEAEGVFEKGSFKTLTLPSLPLNSADEIYELRVTGRTQDEILFSNSTRLSFETKRISVFI 133
Db     66 LEYSGLKKRHLHCISFLVPPPAGGTEEVATIRVSG-VGNNISFEEKKKVLIQRQGNGTFV 124

Qy    134 QTDKALYKPKQEVKFRIVTLFSDFKPYKTSLNIL-IKDPKSNLIQQWLSQQSDLGVISKT 192
Db    125 QTDKPLYTPGQQVYFRIVTMDSNFVPVNDKYSMVELQDPNSNRIAQWLEVVPEQGIVDLS 184

Qy    193 FQLSSHPILGDWSIQVQVNDQTYYQSFQVSEY----------VLPKFEVTLQTPLYCSM 241
Db    185 FQLAPEAMLGTYT--VAVAEGKTFGTFSVEEYVLSPFLLLLSSVLPKFKVEVVEPKELST 242

Qy    242 NSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGKKKNITKTFKINGSANFSFNDEEMKN 301
Db    243 VQESFLVKICCRYTYGKPMLGAVQVSVC------QKAN----------TYWYREVEREQ 285

Qy    302 VMDSSNGLSEYLD----LSSP-------------GPVEILTTVTESVTGISRNVSTNVF 343
Db    286 LPDKCRNLSGQTDKTGCFSAPVDMATFDLIGYAYSHQINIVATVVEEGTGVEANATQNIY 345

Qy    344 FKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVI----------TV 393
Db    346 ISPQMGSMTFEDTSNFYHPNFPFSGKMLLKFPQGGVLPCKNHLVFLVIYGTNGTFNQTLV 405

Qy    394 TQRN-------YTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIE--------FPILEDS 438
Db    406 TDNNGLAPFTLETSGWNGTD---------------VSLEGKFQMEDLVYNPEQVPRYYQN 450

Qy    439 SELQLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVV---------SG 489
Db    451 AYLHLRPFYSTTRSFLGIHRL--------------NGPLKCGQPQEVLVDYYIDPADASP 496

Qy    490 NKRLKELSYMVVSRGQLVAVGKQN----------STMFSLTPENSWTPKACVIVYYIEDD 539
Db    497 DQEI-SFSYYLIGKGSLVMEGQKHLNSKKKGLKASFSLSLTFTSRLAPDPSLVIYAIFPS 555

Qy    540 GEIISDVLKIPVQLVFKNKIKLYWSKVKAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLM 599
Db    556 GGVVADKIQFSVEMCFDN---------QQLPGAEVELQLQAA-PGSLCALRAVDESVLLL 605

Qy    600 NASNDITMENVVHELELYNTGYYLGMF-------------------------------- 626
Db    606 RP-----------DRELSNRSVY-GMFPFWYGHYPYQVAEYDQCPVSGPWDFPQPLIDPM 653

Qy    627 -------------------MNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYA------ 661
Db    654 PQGHSSQRSIIWRPSFSEGTDLFSFFRDVGLKILSNAKIKKP-VDCSHRSPEYSTAMGGG 712

Qy    662 ---ERFMEENEGHIVDIHDFSLGSSPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSIT 718
Db    713 GHPEAFESSTPLHQAE--------DSQVRQYFPETWLWDLFPIGNSGKEAVHVTVPDAIT 764

Qy    719 SWVATGFVISEDLGLGLTTTPVELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDAT 778
Db    765 EWKAMSFCTSQSRGFGLSPT-VGLTAFKPFFVDLTLPYSVVRGESFRLTATIFNYLKDCI 823

Qy    779 EVKVIIEKSDKFDILMTSSEINATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTAL- 837
Db    824 RVQTDLAKSHEYQLESWADS------QTSSCLCADDAKTHHWNITAVKLGHINFTISTKI 877

Qy    838 ---------------SPTASDAVTQMILVKAEGIEKSYSQSILLDLTDNRLQSTLKTLSF 882
Db    878 LDSNEPCGGQKGFVPQKGRSDTLIKPVLVKPEGVLVEKTHSSLLCPKGGKVAS--ESVSL 935

Qy    883 SFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQNMINFAPNIYILDYLTKK 942
Db    936 ELPVDIVPDSTKAYVTVLGDIMGTALQNLDGLVQMPSGCGEQNMVLFAPIIYVLQYLEKA 995

Qy    943 KQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPSGSTWLSAFVLRCFLEADPY 1002
Db    996 GLLTEEIRSRAVGFLEIGYQKELMYKHSNGSYSAFGERDGNGNTWLTAFVTKCFGQAQKF 1055

Qy   1003 IDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGGNKSPVTLTAYIVTSLLGYRKY 1062
Db   1056 IFIDPKNIQDALKWMAGNQLPSGCYANVGNLLHTAMKGGVDDEVSLTAYVTAALLEMGKD 1115

Qy   1063 QPNIDVQESIHFLESEFSRGISDNYTLALITYALSSVGSPKAKE-ALNMLTWRAEQEGGM 1121
Db   1116 VDDPMVSQGLRCLKNS-ATSTTNLYTQALLAYIFSLAGEMDIRNILLKQLDQQAIISGES 1174

Qy   1122 QFWVSSESKLSDS--W-QPRSLDIEVAAYALLSH-----FLQFQTSEGIPIMRWLSRQRN 1173
Db   1175 IYWSQKPTPSSNASPWSEPAAVDVELTAYALLAQLTKPSLTQKEIAKATSIVAWLAKQHN 1234

Qy   1174 SLGGFASTQDTTVALKALSEFA--ALMNTERTNIQVTVT-------GPSSPSPLAVVQ-- 1222
Db   1235 AYGGFSSTQDTVVALQALAKYATTAYMPSEEINLVVKSTENFQRTFNIQSVNRLVFQQDT 1294

Qy   1223 ----PMAVNISANGFGFAICQLNVVYNVKASGSSRRRRSIQNQEAFDLDVAVK----ENK 1274
Db   1295 LPNVPGMYTLEASGQGCVYVQTVLRYNILPP---------TNMKTFSLSVEIGKARCEQP 1345

Qy   1275 DDLNHVDLNVCTSFSGP-GRSGMALMEVNLLSGF--MVPSEAISLSE-TVKKVEYDHGKL 1330
Db   1346 TSPRSLTLTIHTSYVGSRSSSNMAIVEVKMLSGFSPMEGTNQLLLQQPLVKKVEFGTDTL 1405

Qy   1331 NLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEP 1370
Db   1406 NIYLDELIKNTQTYTFTISQSVLVTNLKPATIKVYDYYLP 1445
```

Search completed: January 15, 2004, 18:19:33
Job time : 51 secs

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 14:13:56 ; Search time 268 Seconds
(without alignments)
7060.483 Million cell updates/sec

Title: US-10-020-095-3
Perfect score: 4287
Sequence: 1 atgcagggcccaccgctcct..........ttatggaactttggctgtga 4287

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	110	2.6	4079	4	US-09-016-434-1174	Sequence 1174, Ap
	2	110	2.6	4577	4	US-09-241-606-1	Sequence 1, Appli
c	3	47.4	1.1	7218	1	US-08-232-463-14	Sequence 14, Appl
	4	47.2	1.1	339	4	US-09-311-352B-1	Sequence 1, Appli
	5	46.4	1.1	5211	1	US-08-447-411-1	Sequence 1, Appli
	6	44	1.0	333	4	US-09-241-606-5	Sequence 5, Appli
	7	44	1.0	750	4	US-09-241-606-3	Sequence 3, Appli
	8	41.4	1.0	5398	3	US-09-356-952-11	Sequence 11, Appl
c	9	41.4	1.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
	10	40	0.9	1581	4	US-09-601-198-80	Sequence 80, Appl
c	11	39.4	0.9	15016	4	US-09-601-198-60	Sequence 60, Appl
	12	38.8	0.9	1327	4	US-09-350-756-9	Sequence 9, Appli
c	13	38.8	0.9	3981	2	US-08-955-138-2	Sequence 2, Appli
	14	38.6	0.9	1579	4	US-09-071-035-83	Sequence 83, Appl
	15	38.6	0.9	1680	4	US-09-071-035-81	Sequence 81, Appl
	16	38.2	0.9	3027	4	US-09-328-352-3277	Sequence 3277, Ap
	17	38	0.9	30549	4	US-09-134-001C-322	Sequence 322, App
c	18	37.8	0.9	1062	4	US-09-134-001C-1803	Sequence 1803, Ap
	19	37.8	0.9	28171	4	US-08-961-527-22	Sequence 22, Appl
c	20	37.4	0.9	1152	4	US-09-107-532A-2119	Sequence 2119, Ap
	21	37.2	0.9	526	3	US-08-961-810-14	Sequence 14, Appl
	22	37.2	0.9	526	3	US-08-352-902D-14	Sequence 14, Appl
	23	37.2	0.9	526	4	US-09-265-503B-14	Sequence 14, Appl
	24	37.2	0.9	1707	4	US-09-266-965-68	Sequence 68, Appl
	25	37.2	0.9	53500	4	US-09-266-965-76	Sequence 76, Appl
	26	37	0.9	987	4	US-09-252-991A-9774	Sequence 9774, Ap
c	27	37	0.9	1311	4	US-09-252-991A-10198	Sequence 10198, A
	28	37	0.9	2169	4	US-09-434-408-3	Sequence 3, Appli
c	29	37	0.9	2274	4	US-09-620-312D-154	Sequence 154, App
	30	37	0.9	4248	4	US-09-252-991A-9867	Sequence 9867, Ap
	31	36.4	0.8	1956	4	US-08-559-896B-1	Sequence 1, Appli
	32	36.4	0.8	5924	1	US-08-447-411-44	Sequence 44, Appl
	33	36.4	0.8	5948	2	US-08-662-227-1	Sequence 1, Appli
	34	36.4	0.8	5948	4	US-09-017-947-1	Sequence 1, Appli
	35	36	0.8	55827	4	US-09-813-133A-3	Sequence 3, Appli
c	36	35.8	0.8	426	4	US-09-601-198-41	Sequence 41, Appl
	37	35.8	0.8	549	4	US-09-311-352B-3	Sequence 3, Appli
c	38	35.8	0.8	2478	4	US-09-601-198-52	Sequence 52, Appl
c	39	35.8	0.8	8855	2	US-08-542-003-1	Sequence 1, Appli
c	40	35.8	0.8	8855	2	US-08-322-760A-1	Sequence 1, Appli
c	41	35.8	0.8	8855	4	US-09-236-949-1	Sequence 1, Appli
	42	35.4	0.8	1029	3	US-09-077-675A-4	Sequence 4, Appli
	43	35.4	0.8	1029	4	US-09-077-674-4	Sequence 4, Appli
	44	35.4	0.8	1167	4	US-09-308-003-6	Sequence 6, Appli
	45	35.4	0.8	1497	4	US-09-220-132-94	Sequence 94, Appl

ALIGNMENTS

RESULT 1
US-09-016-434-1174
; Sequence 1174, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1174:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4079 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g177869
US-09-016-434-1174

Query Match 2.6%; Score 110; DB 4; Length 4079;
Best Local Similarity 50.2%; Pred. No. 3.1e-21;

```
 Matches  303;  Conservative    0;  Mismatches  295;  Indels    6;  Gaps    1;

Qy       2649 ATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGAGATGT 2708
Db       2845 ACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCTGTCTCAGTTTTGGGAGACAT 2904

Qy       2709 TCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGTGGTGA 2768
Db       2905 ATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAGATGCCCTATGGCTGTGGAGA 2964

Qy       2769 ACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAAAAGAA 2828
Db       2965 GCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTGGATTATCTAAATGAAACACA 3024

Qy       2829 ACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTACCAGAG 2888
Db       3025 GCAGCTTACTCCAGAGGTCAAGTCCAAGGCCATTGGCTATCTCAACACTGGTTACCAGAG 3084

Qy       2889 AGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGG------AATTATGA 2942
Db       3085 ACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACCTTTGGGGAGCGATATGGCAG 3144

Qy       2943 CCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCC 3002
Db       3145 GAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAGACTTTTGCCCAAGCTCGAGC 3204

Qy       3003 TTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCA 3062
Db       3205 CTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTCATATGGCTCTCCCAGAGGCA 3264

Qy       3063 GAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGG 3122
Db       3265 GAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTCAACAATGCCATAAAGGGAGG 3324

Qy       3123 CAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAA 3182
Db       3325 AGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATCGCCCTTCTGGAGATTCCTCT 3384

Qy       3183 GTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAG 3242
Db       3385 CACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGCCTGGAGTCAGCCTGGAAGAC 3444

Qy       3243 AGGA 3246
Db       3445 AGCA 3448
```

RESULT 2
US-09-241-606-1
; Sequence 1, Application US/09241606
; Patent No. 6472140
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.
; APPLICANT: Kovacs, Dora
; APPLICANT: Saunders, Aleister J.
; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
; TITLE OF INVENTION: Alzheimer's Disease
; FILE REFERENCE: 0609.4460003
; CURRENT APPLICATION NUMBER: US/09/241,606
; CURRENT FILING DATE: 1999-02-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 4577
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: sig_peptide
;   LOCATION: (44)..(112)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (44)..(4468)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (113)..(4468)
US-09-241-606-1

```
  Query Match             2.6%;  Score 110;  DB 4;  Length 4577;
  Best Local Similarity   50.2%;  Pred. No. 3.3e-21;
  Matches  303;  Conservative    0;  Mismatches  295;  Indels    6;  Gaps    1;

Qy       2649 ATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGAGATGT 2708
Db       2845 ACTGCCACCAAATGTGGTAGAAGAATCTGCCCGAGCTTCTGTCTCAGTTTTGGGAGACAT 2904

Qy       2709 TCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGTGGTGA 2768
Db       2905 ATTAGGCTCTGCCATGCAAAACACACAAAATCTTCTCCAGATGCCCTATGGCTGTGGAGA 2964

Qy       2769 ACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAAAAGAA 2828
Db       2965 GCAGAATATGGTCCTCTTTGCTCCTAACATCTATGTACTGGATTATCTAAATGAAACACA 3024

Qy       2829 ACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTACCAGAG 2888
Db       3025 GCAGCTTACTCCAGAGATCAAGTCCAAGGCCATTGGCTATCTCAACACTGGTTACCAGAG 3084

Qy       2889 AGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGG------AATTATGA 2942
Db       3085 ACAGTTGAACTACAAACACTATGATGGCTCCTACAGCACCTTTGGGGAGCGATATGGCAG 3144

Qy       2943 CCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCC 3002
Db       3145 GAACCAGGGCAACACCTGGCTCACAGCCTTTGTTCTGAAGACTTTTGCCCAAGCTCGAGC 3204

Qy       3003 TTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCA 3062
Db       3205 CTACATCTTCATCGATGAAGCACACATTACCCAAGCCCTCATATGGCTCTCCCAGAGGCA 3264

Qy       3063 GAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGG 3122
Db       3265 GAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTCAACAATGCCATAAAGGGAGG 3324

Qy       3123 CAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAA 3182
Db       3325 AGTAGAAGATGAAGTGACCCTCTCCGCCTATATCACCATCGCCCTTCTGGAGATTCCTCT 3384

Qy       3183 GTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAG 3242
Db       3385 CACAGTCACTCACCCTGTTGTCCGCAATGCCCTGTTTTGCCTGGAGTCAGCCTGGAAGAC 3444

Qy       3243 AGGA 3246
Db       3445 AGCA 3448
```

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
;   APPLICANT: DORNER, F.
;   APPLICANT: SCHEIFLINGER, F.
;   APPLICANT: FALKNER, F. G.
;   TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;   NUMBER OF SEQUENCES: 52
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley & Lardner
;   STREET: 1800 Diagonal Road, Suite 500
;   CITY: Alexandria
;   STATE: VA
;   COUNTRY: USA
;   ZIP: 22313-0299
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14

  Query Match             1.1%;  Score 47.4;  DB 1;  Length 7218;
  Best Local Similarity   4.5%;  Pred. No. 0.008;
  Matches   18;  Conservative  217;  Mismatches  168;  Indels    0;  Gaps    0;

Qy       1149 AGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAGTACTGGAG 1208
                | | |||     || ||| | ||       :: :: :: : :::::::  : :  ::: : . ::::
Db       1457 AAAGAGATAGAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398

Qy       1209 CGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTATACTGTCCC 1268
                :::   :: :: :::::  :::::: :::::  :    :::::: :::  : :  :
Db       1397 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1338

Qy       1269 CCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAGCTACAGTT 1328
                 :::: ::::       ::::  :::      ::     ::::::     :: :::  : ::
Db       1337 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1278

Qy       1329 GAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTTAAGTCTCC 1388
              :::::    :          :: :: ::::: :: : :: ::    : ::    :    :::
Db       1277 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1218

Qy       1389 TAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGATCGCCTTT 1448
               :: :::: : : :    ::  ::::: :::::: ::::: : ::::: ::::   :
Db       1217 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1158

Qy       1449 TGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTAGTATCCAG 1508
                :::   :: ::   :: :: :: ::: ::  :::::::  :::  : : :: :: :   ::
Db       1157 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1098

Qy       1509 GGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTC 1551
              :::: ::   :: ::  : ::::::: ::::      |||  | |
Db       1097 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCC 1055

RESULT 4
US-09-311-352B-1
; Sequence 1, Application US/09311352B
; Patent No. 6329500
; GENERAL INFORMATION:
;  APPLICANT:  Webb, Donna J.
;  APPLICANT:  Gonias, Steven L.
;  TITLE OF INVENTION:  Transforming Growth Factor-beta Binding Site
;  FILE REFERENCE:  00370-02
;  CURRENT APPLICATION NUMBER: US/09/311,352B
;  CURRENT FILING DATE:  1999-05-13
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 339
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-311-352B-1

  Query Match             1.1%;  Score 47.2;  DB 4;  Length 339;
  Best Local Similarity   53.0%;  Pred. No. 0.0011;
  Matches  125;  Conservative    0;  Mismatches  108;  Indels    3;  Gaps    1;

Qy       2056 GTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATGGGTTACAGGATT 2115
                || |||||| | || || ||||| ||||| |||                 |    |
Db         97 GTACGAAAGTACTTCCCTGAGACATGGATCTGGGATTTGGTGGTGGTAAACTCAGCAGGG 156

Qy       2116 TACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTT 2175
                     ||  | | |||||| || |||||  | |||||      |||  |||      |  ||
Db        157 GTGGCTGAGGTAGGAGTAACAGTCCCTGACACCATCACCGAGTGGAAGGCAGGGGCCTTC 216

Qy       2176 GTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTC 2235
                     | ||||| ||      || |||||  |   |  | ||| |          |||| |||||||
Db        217 TGCCTGTCTGAAGATGCTGGACTTGGTATCTCTTCCACTGCCTCT---CTCCGAGCCTTC 273

Qy       2236 CAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGAATT 2291
              || || || ||   |        | |    | || |||||||| ||   | || || |   ||
Db        274 CAGCCCTTCTTTGTGGAGCTCACAATGCCTTACTCTGTGATTCGTGGAGAGGCCTT 329

RESULT 5
US-08-447-411-1
; Sequence 1, Application US/08447411
; Patent No. 5773243
;  GENERAL INFORMATION:
;    APPLICANT:  FRITZINGER, DAVID C.
;    APPLICANT:  BREDEHORST, REINHARD
;    APPLICANT:  VOGEL, CARL-WILHELM
;    TITLE OF INVENTION:  DNA ENCODING COBRA C3, CVF1, AND CVF2
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,411
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/043,747
;      FILING DATE:  07-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773243man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  1126-101-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5211 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
```

```
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  9..4961
US-08-447-411-1

  Query Match              1.1%;  Score 46.4;  DB 1;  Length 5211;
  Best Local Similarity   50.4%;  Pred. No. 0.012;
  Matches  113;  Conservative    0;  Mismatches  111;  Indels    0;  Gaps    0;

Qy       3376 TCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCA 3435
              ||| |||      || | ||  ||     |    |   | ||| |||   |||||||||   |
Db       3609 TCAACAGGAAGGAATCGTTGGGAAGAATATAATGCTCGCACCCATAATATTGAAGGCACT 3668

Qy       3436 GCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATG 3495
               |||||||  || |   | |    |    |         |    | ||||||     ||  |  |
Db       3669 TCCTATGCCTTGTTGGCCCTGCTGAAAATGAAGAAATTTGCTGAGGTCGGTCCTGTAGTC 3728

Qy       3496 AGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACT 3555
               || |||||  |        || | |  ||      |||    | || |   || ||  |   ||    |
Db       3729 AGATGGCTGATAGATCAGAAATATTATGGGGAACATATGGACAAACCCAAGCAACAGTT 3788

Qy       3556 GTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATAC 3599
                |||   ||   | |||||  |||||| ||        |   |||   |||
Db       3789 ATGGTGTTTCAAGCTCTTGCTGAATATGAGATTCAGATGCCTAC 3832


RESULT 6
US-09-241-606-5
; Sequence 5, Application US/09241606
; Patent No. 6472140
; GENERAL INFORMATION:
;  APPLICANT: Tanzi, Rudolph E.
;  APPLICANT: Kovacs, Dora
;  APPLICANT: Saunders, Aleister J.
;  TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
;  TITLE OF INVENTION: Alzheimer's Disease
;  FILE REFERENCE: 0609.4460003
;  CURRENT APPLICATION NUMBER: US/09/241,606
;  CURRENT FILING DATE: 1999-02-02
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 333
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(333)
;   OTHER INFORMATION: A  Binding Domain
US-09-241-606-5

  Query Match              1.0%;  Score 44;  DB 4;  Length 333;
  Best Local Similarity   63.0%;  Pred. No. 0.0092;
  Matches   68;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
                | || |    ||   ||| |||| | |   | |||      |||   |   ||| |||||  | ||
Db         19 GCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAGGGCGGTTTCTCCTCC 78

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCC 3588
                || ||||| ||     |||||| |   | ||||||||   ||| || ||||
Db         79 ACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGAGCC 126


RESULT 7
US-09-241-606-3
; Sequence 3, Application US/09241606
; Patent No. 6472140
; GENERAL INFORMATION:
;  APPLICANT: Tanzi, Rudolph E.
;  APPLICANT: Kovacs, Dora
;  APPLICANT: Saunders, Aleister J.
;  TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for
;  TITLE OF INVENTION: Alzheimer's Disease
;  FILE REFERENCE: 0609.4460003
;  CURRENT APPLICATION NUMBER: US/09/241,606
;  CURRENT FILING DATE: 1999-02-02
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;   LENGTH: 750
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(750)
;   OTHER INFORMATION: A /LRP Binding Domain
US-09-241-606-3

  Query Match              1.0%;  Score 44;  DB 4;  Length 750;
  Best Local Similarity   63.0%;  Pred. No. 0.016;
  Matches   68;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
                | || |    ||   ||| |||| | |   | |||      |||   |   ||| |||||  | ||
Db         19 GCAACCAACATCGTGAAGTGGATCACGAAGCAGCAGAATGCCCAGGGCGGTTTCTCCTCC 78

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCC 3588
                || ||||| ||     |||||| |   | ||||||||   ||| || ||||
Db         79 ACCCAGGACACAGTGGTGGCTCTCCATGCTCTGTCCAAATATGGAGCC 126


RESULT 8
US-09-356-952-11
; Sequence 11, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
;  APPLICANT: Boriack-Sjodin, Ann
;  APPLICANT: Margarit, S. M.
;  APPLICANT: Bor-Sogi, Dafna
;  APPLICANT: Cole, Philip
;  APPLICANT: Kuriyan, John
;  TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
;  TITLE OF INVENTION: THEREOF
;  FILE REFERENCE: 600-1-228N
;  CURRENT APPLICATION NUMBER: US/09/356,952
;  CURRENT FILING DATE: 1999-07-19
;  EARLIER APPLICATION NUMBER: 60/093,631
;  EARLIER FILING DATE: 1998-07-21
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 5398
;   TYPE: DNA
;   ORGANISM: Saccharomyces cerevisiae
US-09-356-952-11

  Query Match              1.0%;  Score 41.4;  DB 3;  Length 5398;
  Best Local Similarity   47.2%;  Pred. No. 0.37;
  Matches  126;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy        802 TTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATTACAAAAACATTTAAGATAAAT 861
              |||  ||||||    |   |          | || | |    || ||| || ||           |||
Db       2671 TTTACACCTTTCAACCGTCCTTCTCATAACAGAACCTTTTCAAGAGCAAGAGTTTCAAAA 2730

Qy        862 GGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAAATGTAATGGATTCTTCAAAT 921
                 |         ||| |    | || | ||  ||      |||| |     |  | |  |     ||
Db       2731 AGGAAGAAAAAATATCCATTAACTGTAGACACTTTGAATACAATGAAGAAGAAATCCTCG 2790

Qy        922 GGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCAGTAGAAATTTTAACCACAGTG 981
                 | ||| |||| |    |    ||        || |     || ||    || ||||| |||   | |
Db       2791 CAAATTTTTGAAAAATTAAATAATGCTACAGGTGAACACTTAAAAATTATAAGTAAACCC 2850

Qy        982 ACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACTAATGTGTTCTTCAAGCAACAT 1041
                | |      | |||      | |     || ||||  ||||||  |  |           | || ||   ||
```

```
Db       2851 AAAAGCAGAATTAGGAATTTGGAAATAAATTCAAGCACATACGAACAAATAAATCAGAAT 2910

Qy       1042 GATTACATCATTGAGTTTTTGATTAT 1068
              | ||   |  | ||| |  | ||  ||
Db       2911 GTTTTACTATTGGAGATACTGGAGAAT 2937


RESULT 9
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
;  APPLICANT: Bult et al.
;  TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococ
; Patent No. 6503729
;  TITLE OF INVENTION:  jannaschii
;  FILE REFERENCE: PB275
;  CURRENT APPLICATION NUMBER: US/08/916,421B
;  CURRENT FILING DATE:  1997-08-22
;  PRIOR APPLICATION NUMBER: US 60/024,428
;  PRIOR FILING DATE: 1996-08-22
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 1664976
;   TYPE: DNA
;   ORGANISM: Methanococcus jannaschii
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (28222)..(28222)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (28257)..(28258)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84773)..(84773)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84808)..(84808)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (84812)..(84812)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98120)..(98120)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98159)..(98159)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98239)..(98239)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98266)..(98266)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (98343)..(98343)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (103998)..(103998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (148948)..(148948)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (163385)..(163385)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191989)..(191989)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (191995)..(191995)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (231980)..(231980)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234187)..(234187)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234220)..(234220)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (234814)..(234814)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309398)..(309398)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (309418)..(309418)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312837)..(312837)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (312993)..(312993)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (319226)..(319226)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559167)..(559167)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (559241)..(559241)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (600992)..(600992)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (622708)..(622708)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657081)..(657081)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (657203)..(657203)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (674435)..(674435)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (682442)..(682442)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (713652)..(713652)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (741684)..(741684)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (779455)..(779455)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (779676)..(779676)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (855539)..(855539)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (871619)..(871619)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1084830)..(1084830)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
```

```
;   LOCATION: (1096846)..(1096846)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1119881)..(1119881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1130881)..(1130881)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1310988)..(1310988)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1313224)..(1313224)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349473)..(1349473)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1349491)..(1349491)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1470091)..(1470091)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1569020)..(1569020)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1602912)..(1602912)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1603734)..(1603734)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1637998)..(1637998)
;   OTHER INFORMATION: n equals a, t, c, or g
;   NAME/KEY: misc_feature
;   LOCATION: (1664854)..(1664855)
;   OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

  Query Match             1.0%;  Score 41.4;  DB 4;  Length 1664976;
  Best Local Similarity   52.0%;  Pred. No. 19;
  Matches   93;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy       3725 TTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAG 3784
Db     539968 TTGGAGCTACTGATATGGCTTACATCTATGCAACAGGAGAAACATGGATTAAAGTGCCAA 539909

Qy       3785 AAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATT 3844
Db     539908 AAACAATTAGGGTAGATATAGTTGGAAAAAATGAAAATGTTTCTGCCAAAGATATTGTTT 539849

Qy       3845 TGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTT 3903
Db     539848 TAAGGGTTTGTAAGGAAATTGGGAGAAGAGGAGCAACATACATGGCTATTGAGTATGGT 539790

RESULT 10
US-09-601-198-80
; Sequence 80, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
;  APPLICANT: Cassell, Gail H.
;  APPLICANT:  Chen, Ellson Y.
;  APPLICANT:  Glass, Jennifer S.
;  APPLICANT:  Glass, John I.
;  APPLICANT:  Heiner, Cheryl R.
;  APPLICANT:  Lefkowitz, Elliot
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
;  TITLE OF INVENTION:  UREALYTICUM
;  FILE REFERENCE: UAB-13452/22
;  CURRENT APPLICATION NUMBER: US/09/601,198
;  CURRENT FILING DATE:  2000-12-08
;  PRIOR APPLICATION NUMBER: 60/073,189
;  PRIOR FILING DATE: 1998-01-30
;  NUMBER OF SEQ ID NOS: 181
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
;   LENGTH: 1581
;   TYPE: DNA
;   ORGANISM: Ureaplasma urealyticum
US-09-601-198-80

  Query Match             0.9%;  Score 40;  DB 4;  Length 1581;
  Best Local Similarity   43.8%;  Pred. No. 0.4;
  Matches  175;  Conservative    0;  Mismatches  225;  Indels    0;  Gaps    0;

Qy       3542 CTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAG 3601
Db        578 CTATGGATGCAGTCGATCTTTTAGATACTTGAAAATCGTTTGGATCATTAATTTCATCTA 637

Qy       3602 AAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTGG 3661
Db        638 AAGGTAATTCTGTTGAAAAAAGCGTTTAGCATCTAAATTACAAAATTCATTTGCAAAAA 697

Qy       3662 TACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCA 3721
Db        698 CAGCAGATAAAGTTGTTTACATTTATGATGATGCTATTGTCCCTAATATTTATGATAAGA 757

Qy       3722 ATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATC 3781
Db        758 ATGGCAATCGTAATTTAAAAATTACACAAGAATTTGCTAATTGAATGATTAACCAAAATT 817

Qy       3782 AAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGG 3841
Db        818 ATAAAGGTTATTTAGCTCGTGAATTTTTAAAAATAAATCAAAAAAATGATTTAATTCCTT 877

Qy       3842 ATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAG 3901
Db        878 TACCAATGAGTATTGTTTGACACGCAAGTTATGGTATCATCGGAATGCATCGTATGTTAT 937

Qy       3902 TTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAAT 3941
Db        938 ATACTTTATCAAAAGCTTTTGGAATGCCAAAAGATGAATT 977

RESULT 11
US-09-601-198-60/c
; Sequence 60, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
;  APPLICANT: Cassell, Gail H.
;  APPLICANT:  Chen, Ellson Y.
;  APPLICANT:  Glass, Jennifer S.
;  APPLICANT:  Glass, John I.
;  APPLICANT:  Heiner, Cheryl R.
;  APPLICANT:  Lefkowitz, Elliot
;  TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
;  TITLE OF INVENTION:  UREALYTICUM
;  FILE REFERENCE: UAB-13452/22
;  CURRENT APPLICATION NUMBER: US/09/601,198
;  CURRENT FILING DATE:  2000-12-08
;  PRIOR APPLICATION NUMBER: 60/073,189
;  PRIOR FILING DATE: 1998-01-30
;  NUMBER OF SEQ ID NOS: 181
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
;   LENGTH: 15016
;   TYPE: DNA
;   ORGANISM: Ureaplasma urealyticum
US-09-601-198-60

  Query Match             0.9%;  Score 39.4;  DB 4;  Length 15016;
  Best Local Similarity   48.1%;  Pred. No. 2.9;
  Matches  112;  Conservative    0;  Mismatches  121;  Indels    0;  Gaps    0;
```

```
Qy        703 ACACCATTATATTGTTCTATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTAT 762
Db       2859 ACACCATTTAATGAAGATATTAATAAAAAAGATATTCCAGCAGAAAATATTCCAGTATTT 2800

Qy        763 ACATATGGGAAGCCAGTGAAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGG 822
Db       2799 GAACGTACATACACATTAAATGGAAATTCTAATAGTGCAATTGTACTTTTATCTTTAAAT 2740

Qy        823 GGAAAGAAGAAAAATATTACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTT 882
Db       2739 AATTTAGAGCCTAATATTAAATATAAAATTACTAAAATGTGATTTGAGCAAAAACCAACT 2680

Qy        883 AATGATGAAGAGATGAAAAATGTAATGGATTCTTCAAATGGACTTTCTGAATA 935
Db       2679 AAGGCCTATGCTAATTTAAATTCAAAAGCTAATAACAATGTAATTTATGAAAA 2627
```

RESULT 12
US-09-350-756-9
; Sequence 9, Application US/09350756
; Patent No. 6495143
; GENERAL INFORMATION:
; APPLICANT: U.S. Army Medical Reseach Institute for Infectious Diseases
; APPLICANT: John S. Lee
; APPLICANT: Peter Pushko
; APPLICANT: Michael D. Parker
; APPLICANT: Jonathan F. Smith
; APPLICANT: Mark T. Dertzbaugh
; APPLICANT: Leonard Smith
; TITLE OF INVENTION: Botulinum Neurotoxin Vaccine
; FILE REFERENCE: 003/124/SAP RIID 98-21
; CURRENT APPLICATION NUMBER: US/09/350,756
; CURRENT FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: US 60/092,416
; EARLIER FILING DATE: 1998-07-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 9
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Clostridium botulinum
; FEATURE:
US-09-350-756-9

```
  Query Match             0.9%;  Score 38.8;  DB 4;  Length 1327;
  Best Local Similarity   49.5%;  Pred. No. 0.79;
  Matches  100;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy       1472 AACGATTGAAGGAGTTAAGCTATATGGTAGTATCCAGGGGACAGTTGGTGGCTGTAGGAA 1531
Db        801 AAATAATGTAGGTATTAGAGGTTATATGTATCTTAAAGGGCCTAGAGGTAGCGTAATGAC 860

Qy       1532 AACAAAATTCAACAATGTTCTCTTTAACACCAGAAAATTCTTGGACTCCAAAAGCCTGTG 1591
Db        861 TACAAATGACTACAAACATTTATTTAAATTCAAGTTTGTATAGGGGGACAAAATTTATTA 920

Qy       1592 TAATTGTGTATTATATTGAAGATGATGGGGAAATTATAAGTGATGTTCTAAAAATTCCTG 1651
Db        921 TAAAAAAATATGCTTCTGGAAATAAAGATAATATTGTTAGAAATAATGATCGTGTATATA 980

Qy       1652 TTCAGCTTGTTTTTAAAAATAA 1673
Db        981 TTAATGTAGTAGTTAAAAATAA 1002
```

RESULT 13
US-08-955-138-2/c
; Sequence 2, Application US/08955138A
; Patent No. 5977435
; GENERAL INFORMATION:
; APPLICANT: Lefebvre, Daniel D.
; APPLICANT: Gellatly, Kevin S.
; TITLE OF INVENTION: PLANT PHOSPHATASES
; FILE REFERENCE: PPL97-01
; CURRENT APPLICATION NUMBER: US/08/955,138A
; CURRENT FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3981
; TYPE: DNA
; ORGANISM: SOLANUM TUBEROSUM
US-08-955-138-2

```
  Query Match             0.9%;  Score 38.8;  DB 2;  Length 3981;
  Best Local Similarity   50.0%;  Pred. No. 1.7;
  Matches   97;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy       1503 ATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACACC 1562
Db       3066 ATCTAGACAACAATTGTCTTAAAATTGACACCTTAAATTAATTATCTCAGAGTTAAACTT 3007

Qy       1563 AGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGGGA 1622
Db       3006 AGAACTAACCTAAACGCCACCTACCTAAGGTATATTTGTTATTATGAAGCTAGCAAGTA 2947

Qy       1623 AATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAGCT 1682
Db       2946 AATTATTATTTTGTTTAAACCTTTTTTGAATCATTGAATTAACTAAAAGTGAAATGTT 2887

Qy       1683 ATATTGGAGTAAAG 1696
Db       2886 CAATTGTTGTGAAG 2873
```

RESULT 14
US-09-071-035-83
; Sequence 83, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1579 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-071-035-83

```
  Query Match             0.9%;  Score 38.6;  DB 4;  Length 1579;
  Best Local Similarity   49.3%;  Pred. No. 1;
  Matches  101;  Conservative    0;  Mismatches  104;  Indels    0;  Gaps    0;

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||| | |  | |  |  | ||  |  || | |   |||  |     | | | ||||| | |
Db        188 GTAGCAAAAGCATTGCCGATGATTTCAGAAGATGGAAAAACCTACACGATTTCTTTGAGA 247

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
                ||||     ||||| |    | |    ||| |  |    || || ||||| ||| | ||
Db        248 AAAGAAGCGGTTTGGAGTAACGATGATCCTGTCACAGCACATGATTTGAATATGCTTGG 307

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
               |||  || |  ||||  |  |||      |||| |      |  | |     | | | ||
Db        308 AAAAAAATGATTGATCCTAAAAACGGCTTTGTTTATAGCTTCCTCATCGTTGAAACAATT 367

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTG 1705
               | | || |     || |   ||| |
Db        368 CAAAATGGTGCAGAAATCTCAGCGG 392


RESULT 15
US-09-071-035-81
; Sequence 81, Application US/09071035
; Patent No. 6448043
;  GENERAL INFORMATION:
;    APPLICANT:  Gil H. Choi
;    TITLE OF INVENTION:  Enterococcus faecalis Polynucleotides and Polypeptides
;    NUMBER OF SEQUENCES:  496
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Human Genome Sciences, Inc.
;      STREET:  9410 Key West Avenue
;      CITY:  Rockville
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4Mb storage
;      COMPUTER:  HP Vectra 486/33
;      OPERATING SYSTEM:  MSDOS version 6.2
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/071,035
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  A. Anders Brookes
;      REGISTRATION NUMBER:  36,373
;      REFERENCE/DOCKET NUMBER:  PB369P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (301) 309-8504
;      TELEFAX:  (301) 309-8512
;  INFORMATION FOR SEQ ID NO:  81:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1680 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-09-071-035-81

  Query Match             0.9%;  Score 38.6;  DB 4;  Length 1680;
  Best Local Similarity   49.3%;  Pred. No. 1.1;
  Matches  101;  Conservative    0;  Mismatches  104;  Indels    0;  Gaps    0;

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||| | |  | |  |  | ||  |  || | |   |||  |     | | | ||||| | |
Db        286 GTAGCAAAAGCATTGCCGATGATTTCAGAAGATGGAAAAACCTACACGATTTCTTTGAGA 345

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
                ||||     ||||| |    | |    ||| |  |    || || ||||| ||| | ||
Db        346 AAAGAAGCGGTTTGGAGTAACGATGATCCTGTCACAGCACATGATTTGAATATGCTTGG 405

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
               |||  || |  ||||  |  |||      |||| |      |  | |     | | | ||
Db        406 AAAAAAATGATTGATCCTAAAAACGGCTTTGTTTATAGCTTCCTCATCGTTGAAACAATT 465

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTG 1705
               | | || |     || |   ||| |
Db        466 CAAAATGGTGCAGAAATCTCAGCGG 490

Search completed: January 16, 2004, 21:01:59
Job time : 274 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:          January 16, 2004, 14:38:51 ; Search time 1338 Seconds
                                              (without alignments)
                                              11293.468 Million cell updates/sec

Title:           US-10-020-095-3
Perfect score:   4287
Sequence:        1 atgcagggcccaccgctcct..........ttatggaactttggctgtga 4287

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters:      4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                 1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                 2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                 3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                 4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                 5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                 6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                 7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                 8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                 9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                 10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                 11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                 12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                 13:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
                 14:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                 15:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                 16:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                 17:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                 18:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2584.6	60.3	3033	10	US-09-833-381-1810	Sequence 1810, Ap
	2	1969.4	45.9	2273	12	US-10-108-260A-953	Sequence 953, App
	3	730.8	17.0	1300	13	US-10-133-013-223	Sequence 223, App
	4	541.6	12.6	875	13	US-10-160-162-48	Sequence 48, Appl
	5	541.6	12.6	875	13	US-09-820-649-48	Sequence 48, Appl
	6	272	6.3	354	10	US-09-960-352-12867	Sequence 12867, A
c	7	194.4	4.5	689	13	US-10-027-632-204326	Sequence 204326,
c	8	194.4	4.5	689	14	US-10-027-632-204326	Sequence 204326,
	9	130	3.0	4677	13	US-10-316-253-266	Sequence 266, App
	10	130	3.0	4701	13	US-10-316-253-268	Sequence 268, App
	11	126.8	3.0	4615	10	US-09-880-107-3727	Sequence 3727, Ap
	12	126.8	3.0	4615	13	US-10-006-285-405	Sequence 405, App
	13	125.2	2.9	4488	13	US-09-981-151A-9	Sequence 9, Appli
	14	119.8	2.8	4501	12	US-10-094-886-125	Sequence 125, App
	15	119.4	2.8	1140	9	US-09-764-853-362	Sequence 362, App
	16	119.4	2.8	1140	11	US-09-764-876-17	Sequence 17, Appl
	17	118.6	2.8	2092	11	US-09-764-876-12	Sequence 12, Appl
	18	117.8	2.7	4595	10	US-09-917-800A-1531	Sequence 1531, Ap
	19	117.4	2.7	2092	9	US-09-764-853-163	Sequence 163, App
	20	116.4	2.7	2252	12	US-10-108-260A-1292	Sequence 1292, Ap
	21	114.2	2.7	4527	13	US-09-756-247-3	Sequence 3, Appli
	22	114.2	2.7	4527	13	US-09-756-247-5	Sequence 5, Appli
	23	110	2.6	2041	13	US-10-292-081A-4	Sequence 4, Appli
	24	110	2.6	2041	13	US-10-087-188-1	Sequence 1, Appli
	25	110	2.6	4021	9	US-09-925-301-552	Sequence 552, App
	26	110	2.6	4422	9	US-09-873-403-4	Sequence 4, Appli
	27	110	2.6	4530	13	US-10-292-081A-2	Sequence 2, Appli
	28	110	2.6	4576	13	US-10-292-081A-7	Sequence 7, Appli
	29	110	2.6	4577	9	US-09-873-403-3	Sequence 3, Appli
	30	110	2.6	4577	10	US-09-880-107-2236	Sequence 2236, Ap
	31	110	2.6	4577	12	US-10-331-496A-19	Sequence 19, Appl
	32	110	2.6	4577	13	US-09-873-319-408	Sequence 408, App
	33	110	2.6	4577	13	US-09-960-706-654	Sequence 654, App
	34	110	2.6	4577	13	US-10-292-081A-3	Sequence 3, Appli
	35	110	2.6	4577	13	US-10-292-081A-5	Sequence 5, Appli
	36	110	2.6	4577	14	US-10-052-817-1	Sequence 1, Appli
	37	110	2.6	4577	15	US-10-076-816-56	Sequence 56, Appl
	38	110	2.6	4809	13	US-10-240-965-178	Sequence 178, App
	39	110	2.6	4823	13	US-09-971-392-172	Sequence 172, App
	40	110	2.6	5092	13	US-10-006-285-404	Sequence 404, App
	41	110	2.6	5092	13	US-09-971-429B-2	Sequence 2, Appli
	42	106.4	2.5	4771	13	US-09-756-247-27	Sequence 27, Appl
	43	99	2.3	4426	12	US-10-094-886-123	Sequence 123, App
	44	93.4	2.2	451	11	US-09-918-995-12615	Sequence 12615, A
	45	85.6	2.0	5191	13	US-10-006-285-206	Sequence 206, App

ALIGNMENTS

```
RESULT 1
US-09-833-381-1810
; Sequence 1810, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
;  APPLICANT: Robison, Keith E.
;  TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs
;  FILE REFERENCE: 5800-119
;  CURRENT APPLICATION NUMBER: US/09/833,381
;  CURRENT FILING DATE:  2001-04-11
;  PRIOR APPLICATION NUMBER: 09/516,448
;  PRIOR FILING DATE: 2000-02-29
;  NUMBER OF SEQ ID NOS: 2050
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1810
;   LENGTH: 3033
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-833-381-1810

  Query Match             60.3%;  Score 2584.6;  DB 10;  Length 3033;
  Best Local Similarity   98.0%;  Pred. No. 0;
  Matches 2648;  Conservative    0;  Mismatches    4;  Indels   51;  Gaps    1;

Qy       1636 GTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAGCTATATTGGAGTAAA 1695
              || | |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 GTCCGAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAGCTATATTGGAGTAAA 65

Qy       1696 GTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTGACACAGCCTGACTCC 1755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTGACACAGCCTGACTCC 125

Qy       1756 ATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAATGCCTCTAATGATATT 1815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 ATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAATGCCTCTAATGATATT 185

Qy       1816 ACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGATATTATTTAGGCATG 1875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        186 ACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGATATTATTTAGGCATG 245

Qy       1876 TTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTATTGACAGATGCAAAC 1935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTATTGACAGATGCAAAC 305

Qy       1936 CTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATATGCTGAGAGGTTTATG 1995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 CTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATATGCTGAGAGGTTTATG 365

Qy       1996 GAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTGGGTAGCAGTCCACAT 2055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 GAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTGGGTAGCAGTCCACAT 425

Qy       2056 GTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATGGGTTACAGGATT 2115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 GTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATGGGTTACAGGATT 485

Qy       2116 TACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTT 2175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 TACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTT 545

Qy       2176 GTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTC 2235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 GTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTC 605

Qy       2236 CAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGAATTTGCT 2295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 CAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGAATTTGCT 665

Qy       2296 TTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAGGTAATCATTGAG 2355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 TTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAGGTAATCATTGAG 725

Qy       2356 AAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAATGCCACAGGCCACCAG 2415
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        726 AAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAATGCCACAGGCCACCAG 785

Qy       2416 CAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTTCCCATCAGGCCAACA 2475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 CAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTTCCCATCAGGCCAACA 845

Qy       2476 CATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACTGCTTCTGATGCTGTC 2535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db        846 CATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACTGCTTCTGATGCTATC 905

Qy       2536 ACCCAGATGATTTAGTAAAGGCTGAAGGAATAGAAAAATCATATTCACAATCCATCTTA 2595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 ACCCAGATGATTTAGTAAAGGCTGAAGGAATAGAAAAATCATATTCACAATCCATCTTA 965

Qy       2596 TTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGTTTCTCATTTCCT 2655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        966 TTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGTTTCTCATTTCCT 1025

Qy       2656 CCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGAGATGTTCTTGGT 2715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 CCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGAGATGTTCTTGGT 1085

Qy       2716 CCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGTGGTGAACAGAAC 2775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 CCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGTGGTGAACAGAAC 1145

Qy       2776 ATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAAAAGAAACAACTG 2835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 ATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAAAAGAAACAACTG 1205

Qy       2836 ACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTACCAGAGAGAACTT 2895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 ACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTACCAGAGAGAACTT 1265

Qy       2896 CTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGACCCTTCTGGGAGC 2955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 CTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGACCCTTCTGGGAGC 1325

Qy       2956 ACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCTTACATAGATATT 3015
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 ACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCTTACATAGATATT 1385

Qy       3016 GATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAGAAATCCAACGGT 3075
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1386 GATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAGAAATCCAACGGT 1445

Qy       3076 GAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCA 3135
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 GAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCA 1505

Qy       3136 GTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAAC 3195
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 GTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAAC 1565

Qy       3196 ATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGAC 3255
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 ATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGAC 1625

Qy       3256 AATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAG 3315
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 AATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAG 1685

Qy       3316 GAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTG 3375
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1686 GAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTG 1745

Qy       3376 TCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCA 3435
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1746 TCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCA 1805

Qy       3436 GCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATG 3495
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1806 GCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATG 1865

Qy       3496 AGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACT 3555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1866 AGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACT 1925

Qy       3556 GTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATC 3615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1926 GTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATC 1985

Qy       3616 CAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT------------------------ 3651
              ||||||||||||||||||||||||||||||||||||
Db       1986 CAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTTCTGATTGACACACAC 2045

Qy       3652 ---------------------------CTTGCTGTGGTACAGCCAATGGCAGTTAATATT 3684
                                         |||||||||||||||||||||||||||||||||
Db       2046 AACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCAATGGCAGTTAATATT 2105

Qy       3685 TCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCT 3744
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2106 TCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCT 2165

Qy       3745 TCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTT 3804
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2166 TCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTT 2225

Qy       3805 GCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTT 3864
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2226 GCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTT 2285

Qy       3865 TCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATG 3924
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2286 TCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATG 2345

Qy       3925 GTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGA 3984
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2346 GTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGA 2405
```

```
Qy       3985 AAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCT 4044
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2406 AAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCT 2465

Qy       4045 GTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTAT 4104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2466 GTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTAT 2525

Qy       4105 GAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGAC 4164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2526 GAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGAC 2585

Qy       4165 CTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCAT 4224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2586 CTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCAT 2645

Qy       4225 CACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTG 4284
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2646 CACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTG 2705

Qy       4285 TGA 4287
              |||
Db       2706 TGA 2708
```

RESULT 2
US-10-108-260A-953
; Sequence 953, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 953
; LENGTH: 2273
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-953

```
  Query Match             45.9%;  Score 1969.4;  DB 12;  Length 2273;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1970;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 125

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 185

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 245

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 305

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 365

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 425

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 485

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 545

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 605

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 665

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        666 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAG 725

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 785

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 845

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        846 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 905

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 965

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        966 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1025

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1085

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1145

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1205

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1265

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1325

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1385

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1386 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1445

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1505

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1565
```

```
Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1625

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1685

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1686 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTAAAAATAAGATAAAG 1745

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1746 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1805

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1806 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1865

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1866 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1925

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1926 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1985

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAAT 1971
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1986 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAAT 2036
```

```
RESULT 3
US-10-133-013-223
; Sequence 223, Application US/10133013
; Publication No. US20030166903A1
; GENERAL INFORMATION:
;  APPLICANT: Astromoff, Anna
;  APPLICANT:  Bandman, Olga
;  APPLICANT:  Cocks, Benjamin G.
;  TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE
;  FILE REFERENCE: PA-0049 US
;  CURRENT APPLICATION NUMBER: US/10/133,013
;  CURRENT FILING DATE:  2002-04-25
;  PRIOR APPLICATION NUMBER: 60/287,067
;  PRIOR FILING DATE: 2001-04-27
;  NUMBER OF SEQ ID NOS: 271
;  SOFTWARE: PERL Program
; SEQ ID NO 223
;   LENGTH: 1300
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030166903A1 3658034cB1
US-10-133-013-223

  Query Match             17.0%;  Score 730.8;  DB 13;  Length 1300;
  Best Local Similarity   93.6%;  Pred. No. 4.8e-179;
  Matches  805;  Conservative    0;  Mismatches    2;  Indels   53;  Gaps    2;

Qy       3479 AGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCAT 3538
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db          1 AGGGAATCCCAATTATGAGGCGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCAT 60

Qy       3539 CTACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATA 3598
              ||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||||
Db         61 CTACTCAGGATACCACTGTGGCTTTAAAGG--CTGTCTGAATTTGCAGCCCTAATGAATA 118

Qy       3599 CAGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT------- 3651
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 CAGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGT 178

Qy       3652 --------------------------------------------CTTGCTGTGGTACAGC 3667
                                                          ||||||||||||||||
Db        179 TTCTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGC 238

Qy       3668 CAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTG 3727
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTG 298

Qy       3728 TATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAG 3787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 TATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAG 358

Qy       3788 CCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGA 3847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 CCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGA 418

Qy       3848 ATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACC 3907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 ATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACC 478

Qy       3908 TATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAG 3967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAG 538

Qy       3968 TGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTT 4027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 TGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTT 598

Qy       4028 GTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGT 4087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 GTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGT 658

Qy       4088 CCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGA 4147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGA 718

Qy       4148 AGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAG 4207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 AGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAG 778

Qy       4208 CTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACT 4267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 CTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACT 838

Qy       4268 TTATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||||
Db        839 TTATGGAACTTTGGCTGTGA 858
```

```
RESULT 4
US-10-160-162-48
; Sequence 48, Application US/10160162
; Publication No. US20030166541A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 83 Human Secreted Proteins
;  FILE REFERENCE: PZ012P2
;  CURRENT APPLICATION NUMBER: US/10/160,162
;  CURRENT FILING DATE:  2002-06-04
;  PRIOR APPLICATION NUMBER: 60/295,558
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: 09/236,557
;  PRIOR FILING DATE: 1999-01-26
;  PRIOR APPLICATION NUMBER: PCT/US98/15949
;  PRIOR FILING DATE: 1998-07-29
;  PRIOR APPLICATION NUMBER: 60/054,212
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,209
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,234
;  PRIOR FILING DATE: 1997-07-30
```

```
;  PRIOR APPLICATION NUMBER: 60/054,218
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,214
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,236
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,215
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,211
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,217
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,213
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/055,968
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,969
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/055,972
;  PRIOR FILING DATE: 1997-08-18
;  PRIOR APPLICATION NUMBER: 60/056,561
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,534
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,729
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,543
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,727
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,554
;  PRIOR FILING DATE: 1997-08-19
;  PRIOR APPLICATION NUMBER: 60/056,730
;  PRIOR FILING DATE: 1997-08-19
;  NUMBER OF SEQ ID NOS: 353
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
;   LENGTH: 875
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-160-162-48

  Query Match             12.6%;  Score 541.6;  DB 13;  Length 875;
  Best Local Similarity   99.8%;  Pred. No. 6.4e-130;
  Matches  541;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3746 CTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTG 3805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTG 74

Qy       3806 CTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTT 3865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTT 134

Qy       3866 CGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGG 3925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGG 194

Qy       3926 TGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAA 3985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 TGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAA 254

Qy       3986 AACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTG 4045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 AACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTG 314

Qy       4046 TGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATG 4105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 TGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATG 374

Qy       4106 AGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACC 4165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 AGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACC 434

Qy       4166 TTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATC 4225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATC 494

Qy       4226 ACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGT 4285
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 AMTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGT 554

Qy       4286 GA 4287
              ||
Db        555 GA 556


RESULT 5
US-09-820-649-48
; Sequence 48, Application US/09820649
; Publication No. US20030199683A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: 83 Human Secreted Proteins
;  FILE REFERENCE: PZ012P1
;  CURRENT APPLICATION NUMBER: US/09/820,649
;  CURRENT FILING DATE:  2001-03-30
;  PRIOR APPLICATION NUMBER: US/09/236,557
;  PRIOR FILING DATE: 1999-01-26
;  PRIOR APPLICATION NUMBER: PCT/US98/15949
;  PRIOR FILING DATE: 1998-07-29
;  PRIOR APPLICATION NUMBER: 60/054,212
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,209
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,234
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,218
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,214
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,236
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,215
;  PRIOR FILING DATE: 1997-07-30
;  PRIOR APPLICATION NUMBER: 60/054,211
;  PRIOR FILING DATE: 1997-07-30
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 353
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48
;   LENGTH: 875
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-820-649-48

  Query Match             12.6%;  Score 541.6;  DB 13;  Length 875;
  Best Local Similarity   99.8%;  Pred. No. 6.4e-130;
  Matches  541;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3746 CTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTG 3805
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         15 CTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTG 74

Qy       3806 CTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTT 3865
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         75 CTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTT 134

Qy       3866 CGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGG 3925
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        135 CGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGG 194

Qy       3926 TGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAA 3985
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        195 TGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAA 254
```

```
Qy       3986 AACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTG 4045
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        255 AACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTG 314

Qy       4046 TGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATG 4105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        315 TGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATG 374

Qy       4106 AGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACC 4165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        375 AGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACC 434

Qy       4166 TTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATC 4225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        435 TTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATC 494

Qy       4226 ACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGT 4285
              |:||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        495 AMTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGT 554

Qy       4286 GA 4287
              ||
Db        555 GA 556
```

RESULT 6
US-09-960-352-12867
; Sequence 12867, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12867
;   LENGTH: 354
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (285)
;   OTHER INFORMATION: unsure at all n locations
;   OTHER INFORMATION: Clone ID: 55-LIB3057-014-Q1-K1-F12
US-09-960-352-12867

```
  Query Match             6.3%;  Score 272;  DB 10;  Length 354;
  Best Local Similarity   89.1%;  Pred. No. 5.1e-60;
  Matches  293;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       2084 TTTGGCTAGACACCAACATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTG 2143
              ||.|  | | | ||| | |||||| ||||||||||||||||||  ||||  ||||  ||||
Db         26 TTGGCTTGGCCCCCACCTTGGGTTCCAGGATTTACCAAGAATTCAAAGTTACTGGGCCTG 85

Qy       2144 ATTCTATCACTTCTTGGGTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGAC 2203
              ||||||||||||||||||| || |||| ||||||||||||||||||  | ||||||||||
Db         86 ATTCTATCACTTCTTGGGTTGCAACTGCTTTTGTGATCTCTGAGGATTTAGGTCTTGGAC 145

Qy       2204 TAACAACTACTCCAGTGGAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTC 2263
              |||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 TAACAACTGCTCCAGTGGAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTC 205

Qy       2264 CCTACTCTGTTATCAGAGGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGA 2323
              | |||||||| ||||||||||||||||||||||||||| |||| ||||||||||||||||
Db        206 CATACTCTGTCATCAGAGGTGAAGAATTTGCTTTGGAAGTAACCATATTCAATTATTTGA 265

Qy       2324 AAGATGCCACTGAGGTTAAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGA 2383
              ||||||||||||||||||| ||||||||||||||||||||    |||||||||||||||
Db        266 AAGATGCCACTGAGGTTAANGTAATCATTGAGAAAAGTGATGCTTTTGATATTCTAATGG 325

Qy       2384 CTTCAAGTGAAATAAATGCCACAGGCCAC 2412
              |||| |  |||||||||||| || |||
Db        326 CTTCGAACGAAATAAATGCCACGGGACAC 354
```

RESULT 7
US-10-027-632-204326/c
; Sequence 204326, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204326
;   LENGTH: 689
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-204326

```
  Query Match             4.5%;  Score 194.4;  DB 13;  Length 689;
  Best Local Similarity   99.5%;  Pred. No. 1.3e-39;
  Matches  195;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3352 GAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCA 3411
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        485 GCAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCA 426

Qy       3412 CGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAG 3471
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        425 CGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAG 366

Qy       3472 ACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGT 3531
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        365 ACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGT 306

Qy       3532 TTTGCATCTACTCAGG 3547
              ||||||||||||||||
Db        305 TTTGCATCTACTCAGG 290
```

RESULT 8
US-10-027-632-204326/c
; Sequence 204326, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

```
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 204326
;   LENGTH: 689
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-204326

  Query Match             4.5%;  Score 194.4;  DB 14;  Length 689;
  Best Local Similarity   99.5%;  Pred. No. 1.3e-39;
  Matches  195;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3352 GAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCA 3411

Db        485 GCAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCA 426

Qy       3412 CGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAG 3471

Db        425 CGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAG 366

Qy       3472 ACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGT 3531

Db        365 ACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGT 306

Qy       3532 TTTGCATCTACTCAGG 3547

Db        305 TTTGCATCTACTCAGG 290


RESULT 9
US-10-316-253-266
; Sequence 266, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
;  APPLICANT: The Procter & Gamble Company
;  APPLICANT:  Peters, Kevin
;  APPLICANT:  Thompson, Larry
;  APPLICANT:  Wang, Feng
;  APPLICANT:  Greis, Kenneth
;  TITLE OF INVENTION: Angiogenesis Modulating Proteins
;  FILE REFERENCE: 8865M
;  CURRENT APPLICATION NUMBER: US/10/316,253
;  CURRENT FILING DATE:  2002-12-10
;  PRIOR APPLICATION NUMBER: US 60/355,295
;  PRIOR FILING DATE: 2002-02-08
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 266
;   LENGTH: 4677
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (52)..(4554)
;   OTHER INFORMATION:
US-10-316-253-266

  Query Match             3.0%;  Score 130;  DB 13;  Length 4677;
  Best Local Similarity   47.3%;  Pred. No. 3e-22;
  Matches  575;  Conservative    0;  Mismatches  590;  Indels   51;  Gaps    4;

Qy       2056 GTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATGGGTTACAGGATT 2115

Db       2284 GTGCGGAAGTACTTCCCGGAAACCTGGATCTGGGACATGGTGCCGCTGGACTTATCCGGT 2343

Qy       2116 TACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTT 2175

Db       2344 GACGGTGAGTTGCCGGTAAAGGTCCCTGACACCATCACTGAGTGGAAGGCCAGTGCATTC 2403

Qy       2176 GTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTC 2235

Db       2404 TGCTTGTCTGGAACTACTGGCCTTGGTCTCTCCTCCACCA---TCTCCCATAAAGTCTTC 2460

Qy       2236 CAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGAATTTGCT 2295

Db       2461 CAGCCCTTCTTCCTGGAGCTCACTCTCCCCTACTCTGTGGTTCGAGGCGAAGCATTTATC 2520

Qy       2296 TTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAGGTAATCATTGAG 2355

Db       2521 CTCAAAGCTACCGTACTCAACTACATGCCTCACTGCATTCGGATCCATGTGAGCCTAGAG 2580

Qy       2356 AAA----------------------------AGTGACAAATTTGATATTCTAATGACT 2385

Db       2581 ATGTCTCCTGATTTCCTGGCAGTCCCTGTGGGGAGCCATGAAGATTCTCATTGCATCTGT 2640

Qy       2386 TCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGG 2445

Db       2641 GGAAATGAAAGAAAAACCGTGTCCTGGGCTGTGACCCCAAAGTCGCTGGGGGAGGTGAAC 2700

Qy       2446 GCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACA 2505

Db       2701 TTCACAGCTACTGCAGAAGCTTTGCAGTCTCCAGAACTGTGTGGCAATAAGGTGGCAGAA 2760

Qy       2506 GCTCTTTCACCCACTGCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGA 2565

Db       2761 GTGCCAGCCCTTGTACAGAAGGACACTGTAGTGAAGCCCGTAATAGTCGAGCCTGAAGGA 2820

Qy       2566 ATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGT 2625

Db       2821 ATTGAGAAGGAACAAACGTACAACACACTGTTATGCCCACAAGATGCTGAGTTACA---- 2876

Qy       2626 ACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTT 2685

Db       2877 -----GGAAAATTGGACACTGGATCTTCCAGCCAATGTGGTTGAAGGATCTGCCAGGGCT 2931

Qy       2686 CAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATT 2745

Db       2932 ACACAGTCCGTTTTGGGCGATATACTAGGCTCTGCGATGCAAAACCTCCAGAATCTTCTC 2991

Qy       2746 CGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTGCTCCAAATATTTACATT 2805

Db       2992 CAGATGCCCTATGGCTGTGGGGAACAAAACATGGTCCTCTTCGTCCCTAACATCTACGTT 3051

Qy       2806 TTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCA 2865

Db       3052 CTGGAGTATCTCAATGAGACACAGCAGCTGACAGAGGCGATCAAGTCCAAGGCCATTAGC 3111

Qy       2866 TTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGT 2925

Db       3112 TACCTCATCAGCGGGTACCAGAGGCAGCTGAACTATCAGCACAGTGACGGTTCATACAGC 3171

Qy       2926 GCTTTTGGGAATTATG---------ACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTT 2976

Db       3172 ACATTCGGGGACCGCGGTATGAGGCACAGTCAGGGAAACACTTGGCTCACTGCATTTGTG 3231

Qy       2977 TTAAGATGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGA 3036

Db       3232 CTCAAGGCCTTCGCTCAAGCTCAGTCATACATCTATATAGAAAAGACACACATCACAAAT 3291

Qy       3037 ACATACACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGA 3096

Db       3292 GCTTTCAATTGGCTCTCGATGAAACAAAGGGAGAACGGTTGTTTCCAACAGTCTGGATCC 3351

Qy       3097 GTGATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATT 3156
```

```
Db       3352 CTGCTCAACAATGCGATGAAGGGTGGTGTGGATGATGAAGTGACACTCTCTGCCTATATC 3411

Qy       3157 GTAACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATC 3216

Db       3412 ACCATTGCTCTGCTGGAGATGCCCCTGCCTGTCACTCACAGTGTTGTTCGTAATGCTCTA 3471

Qy       3217 CATTTTTTGGAGTCTG 3232

Db       3472 TTCTGCCTGGAAACGG 3487


RESULT 10
US-10-316-253-268
; Sequence 268, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
;  APPLICANT: The Procter & Gamble Company
;  APPLICANT:  Peters, Kevin
;  APPLICANT:  Thompson, Larry
;  APPLICANT:  Wang, Feng
;  APPLICANT:  Greis, Kenneth
;  TITLE OF INVENTION: Angiogenesis Modulating Proteins
;  FILE REFERENCE: 8865M
;  CURRENT APPLICATION NUMBER: US/10/316,253
;  CURRENT FILING DATE:  2002-12-10
;  PRIOR APPLICATION NUMBER: US 60/355,295
;  PRIOR FILING DATE: 2002-02-08
;  NUMBER OF SEQ ID NOS: 308
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 268
;   LENGTH: 4701
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (95)..(4597)
;   OTHER INFORMATION:
US-10-316-253-268

  Query Match             3.0%;  Score 130;  DB 13;  Length 4701;
  Best Local Similarity   47.3%;  Pred. No. 3e-22;
  Matches  575;  Conservative    0;  Mismatches  590;  Indels   51;  Gaps    4;

Qy       2056 GTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATGGGTTACAGGATT 2115

Db       2327 GTGCGGAAGTACTTCCCGGAAACCTGGATCTGGGACATGGTGCCGCTGGACTTATCCGGT 2386

Qy       2116 TACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTT 2175

Db       2387 GACGGTGAGTTGCCGGTAAAGGTCCCTGACACCATCACTGAGTGGAAGGCCAGTGCATTC 2446

Qy       2176 GTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTC 2235

Db       2447 TGCTTGTCTGGAACTACTGGCCTTGGTCTCTCCTCCACCA---TCTCCCATAAAGTCTTC 2503

Qy       2236 CAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGAATTTGCT 2295

Db       2504 CAGCCCTTCTTCCTGGAGCTCACTCTCCCCTACTCTGTGGTTCGAGGCGAAGCATTTATC 2563

Qy       2296 TTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAGGTAATCATTGAG 2355

Db       2564 CTCAAAGCTACCGTACTCAACTACATGCCTCACTGCATTCGGATCCATGTGAGCCTAGAG 2623

Qy       2356 AAA-----------------------------AGTGACAAATTTGATATTCTAATGACT 2385

Db       2624 ATGTCTCCTGATTTCCTGGCAGTCCCTGTGGGGAGCCATGAAGATTCTCATTGCATCTGT 2683

Qy       2386 TCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGG 2445

Db       2684 GGAAATGAAAGAAAAACCGTGTCCTGGGCTGTGACCCCAAAGTCGCTGGGGGAGGTGAAC 2743

Qy       2446 GCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACA 2505

Db       2744 TTCACAGCTACTGCAGAAGCTTTGCAGTCTCCAGAACTGTGTGGCAATAAGGTGGCAGAA 2803

Qy       2506 GCTCTTTCACCCACTGCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGA 2565

Db       2804 GTGCCAGCCCTTGTACAGAAGGACACTGTAGTGAAGCCCGTAATAGTCGAGCCTGAAGGA 2863

Qy       2566 ATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGT 2625

Db       2864 ATTGAGAAGGAACAAACGTACAACACACTGTTATGCCCACAAGATGCTGAGTTACA---- 2919

Qy       2626 ACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTT 2685

Db       2920 -----GGAAAATTGGACACTGGATCTTCCAGCCAATGTGGTTGAAGGATCTGCCAGGGCT 2974

Qy       2686 CAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATT 2745

Db       2975 ACACAGTCCGTTTTGGGCGATATACTAGGCTCTGCGATGCAAAACCTCCAGAATCTTCTC 3034

Qy       2746 CGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATT 2805

Db       3035 CAGATGCCCTATGGCTGTGGGGAACAAAACATGGTCCTCTTCGTCCCTAACATCTACGTT 3094

Qy       2806 TTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCA 2865

Db       3095 CTGGAGTATCTCAATGAGACACAGCAGCTGACAGAGGCGATCAAGTCCAAGGCCATTAGC 3154

Qy       2866 TTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGT 2925

Db       3155 TACCTCATCAGCGGGTACCAGAGGCAGCTGAACTATCAGCACAGTGACGGTTCATACAGC 3214

Qy       2926 GCTTTTGGGAATTATG---------ACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTT 2976

Db       3215 ACATTCGGGGACCGCGGTATGAGGCACAGTCAGGGAAACACTTGGCTCACTGCATTTGTG 3274

Qy       2977 TTAAGATGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGA 3036

Db       3275 CTCAAGGCCTTCGCTCAAGCTCAGTCATACATCTATATAGAAAAGACACACATCACAAAT 3334

Qy       3037 ACATACACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGA 3096

Db       3335 GCTTTCAATTGGCTCTCGATGAAACAAAGGGAGAACGGTTGTTTCCAACAGTCTGGATCC 3394

Qy       3097 GTGATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATT 3156

Db       3395 CTGCTCAACAATGCGATGAAGGGTGGTGTGGATGATGAAGTGACACTCTCTGCCTATATC 3454

Qy       3157 GTAACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATC 3216

Db       3455 ACCATTGCTCTGCTGGAGATGCCCCTGCCTGTCACTCACAGTGTTGTTCGTAATGCTCTA 3514

Qy       3217 CATTTTTTGGAGTCTG 3232

Db       3515 TTCTGCCTGGAAACGG 3530


RESULT 11
US-09-880-107-3727
; Sequence 3727, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
```

```
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3727
;   LENGTH: 4615
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 X54380
US-09-880-107-3727

  Query Match             3.0%;  Score 126.8;  DB 10;  Length 4615;
  Best Local Similarity   51.6%;  Pred. No. 2e-21;
  Matches  318;  Conservative    0;  Mismatches  292;  Indels    6;  Gaps    1;

Qy       2638 TTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCA 2697
Db       2838 TTGTCCTTGAAGCTCCCATCAAATGTGGTCAAAGAATCTGCCAGAGCTTCTTTCTCAGTT 2897

Qy       2698 ATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTAT 2757
Db       2898 CTGGGTGACATATTAGGTTCTGCTATGCAAAATATACAAAATCTCCTCCAGATGCCATAT 2957

Qy       2758 GGCTGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTG 2817
Db       2958 GGCTGTGGAGAACAGAACATGGTCCTATTTGCTCCTAACATCTATGTCTTGAACTATCTG 3017

Qy       2818 ACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAA 2877
Db       3018 AATGAAACCCAGCAGCTGACGCAGGAGATCAAGGCCAAGGCCGTTGGCTATCTCATCACT 3077

Qy       2878 GGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGG--- 2934
Db       3078 GGTTACCAGAGACAGCTGAACTACAAACACCAAGATGGCTCCTACAGCACCTTTGGGGAA 3137

Qy       2935 ---AATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTT 2991
Db       3138 CGATATGGCAGGAACCAGGGCAACACTTGGCTCACAGCTTTTGTACTGAAGACTTTCGCC 3197

Qy       2992 GAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTT 3051
Db       3198 CAGGCTCGATCCTACATCTTCATTGATGAAGCACACATTACCCAATCTCTCACGTGGCTC 3257

Qy       3052 AAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAG 3111
Db       3258 TCCCAGATGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTCAACAATGCC 3317

Qy       3112 CTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTG 3171
Db       3318 ATAAAGGGAGGTGTAGAAGATGAAGCGACCCTCTCCGCCTATGTTACTATTGCCCTTCTG 3377

Qy       3172 GGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCT 3231
Db       3378 GAAATTCCTCTCCCAGTCACTAACCCTATTGTTCGCAATGCCCTGTTCTGCCTGGAGTCA 3437

Qy       3232 GAATTCAGTAGAGGAA 3247
Db       3438 GCCTGGAATGTAGCAA 3453

RESULT 12
US-10-006-285-405
; Sequence 405, Application US/10006285
; Publication No. US20030165854A1
; GENERAL INFORMATION:
;  APPLICANT: Mary Jane Cunningham
;  APPLICANT:  Matthew R. Kaser
;  TITLE OF INVENTION: MARKER GENES RESPONDING TO TREATMENT WITH TOXINS
;  FILE REFERENCE: PA-0039 US
;  CURRENT APPLICATION NUMBER: US/10/006,285
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 514
;  SOFTWARE: PERL Program
; SEQ ID NO 405
;   LENGTH: 4615
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20030165854A1 369213.39
US-10-006-285-405

  Query Match             3.0%;  Score 126.8;  DB 13;  Length 4615;
  Best Local Similarity   51.6%;  Pred. No. 2e-21;
  Matches  318;  Conservative    0;  Mismatches  292;  Indels    6;  Gaps    1;

Qy       2638 TTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCA 2697
Db       2838 TTGTCCTTGAAGCTCCCATCAAATGTGGTCAAAGAATCTGCCAGAGCTTCTTTCTCAGTT 2897

Qy       2698 ATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTAT 2757
Db       2898 CTGGGTGACATATTAGGTTCTGCTATGCAAAATATACAAAATCTCCTCCAGATGCCATAT 2957

Qy       2758 GGCTGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTG 2817
Db       2958 GGCTGTGGAGAACAGAACATGGTCCTATTTGCTCCTAACATCTATGTCTTGAACTATCTG 3017

Qy       2818 ACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAA 2877
Db       3018 AATGAAACCCAGCAGCTGACGCAGGAGATCAAGGCCAAGGCCGTTGGCTATCTCATCACT 3077

Qy       2878 GGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGG--- 2934
Db       3078 GGTTACCAGAGACAGCTGAACTACAAACACCAAGATGGCTCCTACAGCACCTTTGGGGAA 3137

Qy       2935 ---AATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTT 2991
Db       3138 CGATATGGCAGGAACCAGGGCAACACTTGGCTCACAGCTTTTGTACTGAAGACTTTCGCC 3197

Qy       2992 GAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTT 3051
Db       3198 CAGGCTCGATCCTACATCTTCATTGATGAAGCACACATTACCCAATCTCTCACGTGGCTC 3257

Qy       3052 AAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAG 3111
Db       3258 TCCCAGATGCAGAAGGACAATGGCTGTTTCAGGAGCTCTGGGTCACTGCTCAACAATGCC 3317

Qy       3112 CTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTG 3171
Db       3318 ATAAAGGGAGGTGTAGAAGATGAAGCGACCCTCTCCGCCTATGTTACTATTGCCCTTCTG 3377

Qy       3172 GGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCT 3231
Db       3378 GAAATTCCTCTCCCAGTCACTAACCCTATTGTTCGCAATGCCCTGTTCTGCCTGGAGTCA 3437

Qy       3232 GAATTCAGTAGAGGAA 3247
Db       3438 GCCTGGAATGTAGCAA 3453

RESULT 13
US-09-981-151A-9
; Sequence 9, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
;  APPLICANT: Edinger, Shlomit R
;  APPLICANT:  Gerlach, Valerie
;  APPLICANT:  MacDougall, John R
;  APPLICANT:  Malyankar, Muriel M
;  APPLICANT:  Smithson, Glennda
;  APPLICANT:  Millet, Isabelle
;  APPLICANT:  Peyman, John A
;  APPLICANT:  Stone, David J
;  APPLICANT:  Gunther, Erik
;  APPLICANT:  Ellerman, Karen
;  APPLICANT:  Shimkets, Richard A
```

```
; APPLICANT:  Padigaru, Muralidhara
; APPLICANT:  Guo, Xiaojia
; APPLICANT:  Patturajan, Meera
; APPLICANT:  Taupier Jr, Raymond J
; APPLICANT:  Burgess, Catherine E
; APPLICANT:  Zerhusen, Bryan D
; APPLICANT:  Kekuda, Ramesh
; APPLICANT:  Spytek, Kimberly A
; APPLICANT:  Gangolli, Esha A
; APPLICANT:  Fernandes, Elma R
; APPLICANT:  Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE:  2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 4488
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-981-151A-9

  Query Match             2.9%;  Score 125.2;  DB 13;  Length 4488;
  Best Local Similarity   45.6%;  Pred. No. 5.2e-21;
  Matches  549;  Conservative    0;  Mismatches  628;  Indels   27;  Gaps    2;

Qy       2049 TCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATGGGTTA 2108
Db       2187 TTCTCAGGTCCGCCAGTACTTCCCAGAGACCTGGCTCTGGGATCTGTTTCCTATTGGTAA 2246

Qy       2109 CAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTGGCTAC 2168
Db       2247 CTCGGGGAAGGAGGCGGTCCACGTCACAGTTCCTGACGCCATCACCGAGTGGAAGGCGAT 2306

Qy       2169 TGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAGCTCCA 2228
Db       2307 GAGTTTCTGCACTTCCCAGTCAAGAGGCTTCGGGCTTTCACCCACT---GTTGGACTAAC 2363

Qy       2229 AGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGA 2288
Db       2364 TGCTTTCAAGCCGTTCTTTGTTGACCTGACTCTCCCTTACTCAGTAGTCCGTGGGGAATC 2423

Qy       2289 ATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAGGTAAT 2348
Db       2424 CTTTCGTCTTACTGCCACCATCTTCAATTACCTAAAGGATTGCATCAGGGTTCAGACTGA 2483

Qy       2349 CATTGAGAAAA-----------------------GTGACAAATTTGATATTCTAATGAC 2384
Db       2484 CCTGGCTAAATCGCATGAGTACCAGCTAGAATCATGGGCAGATTCTCAGACCTCCAGTTG 2543

Qy       2385 TTCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGG 2444
Db       2544 TCTCTGTGCTGATGACGCAAAAACCCACCACTGGAACATCACAGCTGTCAAATTGGGTCA 2603

Qy       2445 GGCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCAC 2504
Db       2604 CATTAACTTTACTATTAGTACAAAGATTCTGGACAGCAATGAACCATGTGGGGGCCAGAA 2663

Qy       2505 AGCTCTTTCACCCACTGCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGG 2564
Db       2664 GGGGTTTGTTCCCCAAAAGGGCCGAAGTGACACGCTCATCAAGCCAGTTCTCGTCAAACC 2723

Qy       2565 AATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAG 2624
Db       2724 TGAGGGAGTCCTGGTGGAGAAGACACACAGCTCATTGCTGTGCCCAAAAGGAGGAAAGGT 2783

Qy       2625 TACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGT 2684
Db       2784 GGCATCTGAATCTGTCTCCCTGGAGCTCCCAGTGGACATTGTTCCTGACTCGACCAAGGC 2843

Qy       2685 TCAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGAT 2744
Db       2844 TTATGTTACGGTTCTGGGAGACATTATGGGCACAGCCCTGCAGAACCTGGATGGTCTGGT 2903

Qy       2745 TCGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACAT 2804
Db       2904 GCAGATGCCCAGTGGCTGTGGCGAGCAGAACATGGTCTTGTTTGCTCCCATCATCTATGT 2963

Qy       2805 TTTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTC 2864
Db       2964 CTTGCAGTACCTGGAGAAGGCAGGGCTGCTGACGGAGGAGATCAGGTCTCGGGCAGTGGG 3023

Qy       2865 ATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAG 2924
Db       3024 TTTCCTGGAAATAGGGTACCAGAAGGAGCTGATGTACAAACACAGCAATGGCTCATACAG 3083

Qy       2925 TGCTTTTGGGAATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATG 2984
Db       3084 TGCCTTTGGGGAGCGAGATGGAAATGGAAACACATGGCTGACAGCGTTTGTCACAAAATG 3143

Qy       2985 TTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACAC 3044
Db       3144 CTTTGGCCAAGCTCAGAAATTCATCTTCATTGATCCCAAGAACATCCAGGATGCTCTCAA 3203

Qy       3045 TTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCA 3104
Db       3204 GTGGATGGCAGGAAACCAGCTCCCCAGTGGCTGCTATGCCAACGTGGGAAATCTCCTTCA 3263

Qy       3105 TAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTC 3164
Db       3264 CACAGCTATGAAGGGTGGTGTTGATGATGAGGTCTCCTTGACTGCGTATGTCACAGCTGC 3323

Qy       3165 TCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTT 3224
Db       3324 ATTGCTGGAGATGGGAAAGGATGTAGATGACCCAATGGTGAGTCAGGGTCTACGGTGTCT 3383

Qy       3225 GGAG 3228
Db       3384 CAAG 3387

RESULT 14
US-10-094-886-125
; Sequence 125, Application US/10094886
; Publication No. US20040002120A1
; GENERAL INFORMATION:
;  APPLICANT: Kekuda, Ramesh
;  APPLICANT:  Tchernev, Velizar T.
;  APPLICANT:  Liu, Xiaohong
;  APPLICANT:  Spytek, Kimberly A.
;  APPLICANT:  Patturajan, Meera
;  APPLICANT:  Burgess, Catherine
;  APPLICANT:  Vernet, Corine A.
;  APPLICANT:  Li, Li
```

```
; APPLICANT:  Gorman, Linda
; APPLICANT:  Malyankar, Uriel M.
; APPLICANT:  Boldog, Ferenc
; APPLICANT:  Guo, Xiaojia
; APPLICANT:  Shenoy, Suresh
; APPLICANT:  Padigaru, Muralidhara
; APPLICANT:  Taupier, Raymond J., Jr.
; APPLICANT:  Miller, Charles
; APPLICANT:  Casman, Stacie
; APPLICANT:  Pena, Carol
; APPLICANT:  Gangolli, Esha
; APPLICANT:  Gusev, Vladimir
; APPLICANT:  Smithson, Glennda
; APPLICANT:  Zerhusen, Bryan
; APPLICANT:  Gerlach, Valerie
; APPLICANT:  Pochart, Pascal
; APPLICANT:  Fernandes, Elma
; APPLICANT:  Shimkets, Richard
; APPLICANT:  Rastelli, Luca
; APPLICANT:  Spaderna, Steven
; APPLICANT:  LaRochelle, William
; APPLICANT:  Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-290 B
; CURRENT APPLICATION NUMBER: US/10/094,886
; CURRENT FILING DATE:  2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,182
; PRIOR FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: 60/288,052
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/318,510
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/314,018
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/296,693
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/313,626
; PRIOR FILING DATE: 2001-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 298
; SOFTWARE: PatentIn 2.1
; SEQ ID NO 125
;   LENGTH: 4501
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(4309)
US-10-094-886-125

  Query Match             2.8%;  Score 119.8;  DB 12;  Length 4501;
  Best Local Similarity   45.7%;  Pred. No. 1.3e-19;
  Matches  596;  Conservative    0;  Mismatches  682;  Indels   27;  Gaps    4;

Qy       2040 GGGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAA 2099
Db       2088 GGCTATAATTCAAACAGTAAGAACAAACTTCCCAGAGACATGGATGTGGGACCTCGTCAG 2147

Qy       2100 CATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTG 2159
Db       2148 TGTCGATTCCTCAGGCTCTGCCAATCTTTCGTTCCTCATTCCTGATACGATAACCCAATG 2207

Qy       2160 GGTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGT 2219
Db       2208 GGAGGCAAGTGGCTTTTGTGTGAATGGTGACGTTGGATTTGGCATTTCCTCTACAACCAC 2267

Qy       2220 GGAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAG 2279
Db       2268 T---CTAGAAGTCTCCCAACCTTTCTTTATTGAGATTGCCTCACCCTTTTCGGTTGTTCA 2324

Qy       2280 AGGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGT 2339
Db       2325 AAATGAACAATTTGATTTGATTGTCAATGTCTTCAGCTACCGGAATACATGTGTAGAGAT 2384

Qy       2340 TAAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAA 2399
Db       2385 TTCTGTTCAAGTGGAGGAGTCTCAGAATTATGAAGCAAATATTCATACCTTGAAAATCAA 2444

Qy       2400 TGCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTT 2459
Db       2445 TGGCAGTGAGGTTATTCAAGCTGGAGGGAGGAAAACAAACGTCTGGACTATTATACCTAA 2504

Qy       2460 TCCCATCAGGCCAACACATCTGGGAGAAATTCCT-ATCACAGTCACAGCTCTTTCACCCA 2518
Db       2505 GAAATTGGGTAAAGTGAATATCACTGTAGTTGCTGAGTCCAAACAAAGCAGTGCTTGCCC 2564

Qy       2519 CTGCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGGCTGAAGGAATAGAAAAATCAT 2578
Db       2565 AAATGAAGGAATGGAGCAGCAAAAGCTAAACTGGAAAGACACTGTGGTCCAAAGCTTCTT 2624

Qy       2579 ATTCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTT 2638
Db       2625 AGTAGAGCCTGAAGGTATTGAAAAGGAAAGGACCCAGAGTTTCCTTATCTGTACAGAAGG 2684

Qy       2639 TG-------------------AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGA 2678
Db       2685 TGCCAAAGCCTCCAAGCAGGGAGTTTTGGACTTGCCAAACGATGTAGTAGAAGGGTCAGC 2744

Qy       2679 AAGAGTTCAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTC 2738
Db       2745 CAGAGGCTTTTTCACTGTTGTGGGGGATATTCTAGGACTTGCCTTGCAGAATCTGG---T 2801

Qy       2739 ATTGATTCGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCAAATAT 2798
Db       2802 TGTTCTCCAAATGCCCTATGGAAGTGGAGAGCAGAATGCTGCCCTACTAGCATCTGATAC 2861

Qy       2799 TTACATTTTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGC 2858
Db       2862 TTATGTTCTGGACTATCTGAAATCTACTGAGCAACTGACAGAGGAAGTTCAATCTAAGGC 2921

Qy       2859 TCTTTCATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTC 2918
Db       2922 TTTCTTTCTCTTATCTAATGGTTATCAAAGGCAATTATCTTTCAAAAACTCTGATGGTTC 2981

Qy       2919 TTTCAGTGCTTTTGGGAATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTT 2978
Db       2982 CTATAGTGTGTTTTGGCAGCAGAGTCAGAAAGGAAGCATATGGCTCAGTGCTCTTACTTT 3041

Qy       2979 AAGATGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGAAC 3038
Db       3042 TAAGACATTGGAGAGAATGAAAAAATATGTATTCATTGATGAAAATGTTCAAAAACAGAC 3101

Qy       3039 ATACACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGT 3098
Db       3102 CTTAATCTGGCTTTCAAGCCAACAGAAAACAAGCGGCTGCTTTAAGAATGATGGCCAGCT 3161

Qy       3099 GATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGT 3158
Db       3162 TTTCAACCACGCCTGGGAGGGTGGAGATGAAGAGGACATTTCACTCACTGCGTATGTTGT 3221

Qy       3159 AACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCA 3218
Db       3222 TGGGATGTTCTTTGAAGCTGGGCTCAATTTCACTTTTCCTGCTCTACGAAACGCACTCTT 3281

Qy       3219 TTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCTAGCCCTTATAAC 3278
Db       3282 TTGCCTTGAAGCGGCATTGGACAGTGGTGTCACTAATGGCTATAATCATGCAATTCTAGC 3341
```

```
Qy       3279 TTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTT 3323
              |||||| ||  | | ||  || |      ||       |||| ||||
Db       3342 TTATGCTTTTGCCTTAGCTGGAAAAGAGAAGCAAGTGGAATCTTT 3386

RESULT 15
US-09-764-853-362
; Sequence 362, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PJZ06
;  CURRENT APPLICATION NUMBER: US/09/764,853
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 939
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 362
;   LENGTH: 1140
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1018)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (1119)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-764-853-362

  Query Match             2.8%;  Score 119.4;  DB 9;  Length 1140;
  Best Local Similarity   50.1%;  Pred. No. 6.3e-20;
  Matches  285;  Conservative    4;  Mismatches  280;  Indels    0;  Gaps    0;

Qy       2660 ATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGAGATGTTCTTGGTCCTT 2719
              | |  ||   ||| |      |  | |  |   | ||  |     | |||||  || | ||  |
Db         29 ACATTGTTCCTGACTCGACCAAGGCTTATGTTACGGTTCTGGGAGACATTATGGGCACAG 88

Qy       2720 CCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGTGGTGAACAGAACATGA 2779
              || |  |     | | |       || |  ||||||   |||||||| || |||||||||
Db         89 CCCTGCAGAACCTGGATGGTCTGGTGCAGATGCCCAGTGGCTGTGGCGAGCAGAACATGG 148

Qy       2780 TAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAAAAGAAACAACTGACAG 2839
              |     |||||||| |  || ||  |  ||| | || |||    ||        |  ||||| |
Db        149 TCTTGTTTGCTCCCATCATCTATGTCTTGCAGTACCTGGAGAAGGCAGGGCTGCTGACGG 208

Qy       2840 ATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCT 2899
              |  |  | |           || |      ||  ||     ||| |||:|||  || ||  | |
Db        209 AGGAGATCAGGTCTCGGGCAGTGGGTTTCCTGGAAATAGGGTACSAGAAGGAGCTGATGT 268

Qy       2900 ATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGACCCTTCTGGGAGCACTT 2959
              |  |         ||||||| | |||||| |||||| |    ||        ||| | ||| |
Db        269 ACAAACACAGCAATGGCTCATACAGTGCCTTTGGGGAGCGAGATGGAAATGGAAACACAT 328

Qy       2960 GGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCTTACATAGATATTGATC 3019
              || || |||| |||||    || ||| ||      ||||  |     | |||      |||||||
Db        329 GGCTGACAGCGTTTGTCACAAAATGCTTTGGCCAAGCTCAGAAATTCATCTTCATTGATC 388

Qy       3020 AGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAAT 3079
                ||       | ||      |   ||  |||  |     |:|| | |||       |||  ||    |
Db        389 CCAAGAACATCCAGGATGCTCTCAAGTGGATGGCARGAAACCAGCTCCCCAGTGGCTGCT 448

Qy       3080 TTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAA 3139
               |      |   :||||   |   ||||  |    |     |   |  |||||    ||  |     :||
Db        449 ATGCCAACGTRGGAAATCTCCTTCACACAGCTATGAAGGGTGGTGTTGATGATGARGTCT 508

Qy       3140 CACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTG 3199
              |  |  || || |||  |    |  || |   | ||||        ||||| ||        | ||
Db        509 CCTTGACTGCGTATGTCACAGCTGCATTGCTGGAGATGGGAAAGGATGTAGATGACCCAA 568

Qy       3200 ATGTGCAAGAGTCTATCCATTTTTTGGAG 3228
                |||     ||  |  |  |   |  |  |   ||
Db        569 TGGTGAGTCAGGGTCTACGGTGTCTCAAG 597

Search completed: January 17, 2004, 00:35:05
Job time : 1346 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:          January 16, 2004, 14:02:51 ; Search time 7724 Seconds
                                              (without alignments)
                                              13489.551 Million cell updates/sec

Title:           US-10-020-095-3
Perfect score:   4287
Sequence:        1 atgcagggcccaccgctcct..........ttatggaactttggctgtga 4287

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters:      45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  em_gss_hum:*
                 18:  em_gss_inv:*
                 19:  em_gss_pln:*
                 20:  em_gss_vrt:*
                 21:  em_gss_fun:*
                 22:  em_gss_mam:*
                 23:  em_gss_mus:*
                 24:  em_gss_pro:*
                 25:  em_gss_rod:*
                 26:  em_gss_phg:*
                 27:  em_gss_vrl:*
                 28:  gb_gss1:*
                 29:  gb_gss2:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                     %
Result               Query
   No.     Score     Match Length DB  ID                        Description
------------------------------------------------------------------------------
     1    2708.8     63.2   4485  11  AK029247                  AK029247 Mus muscu
     2    2668.8     62.3   4725  11  AK036799                  AK036799 Mus muscu
     3       893     20.8   1201  9   AL554092                  AL554092 AL554092
c    4     872.8     20.4    941  13  BX391819                  BX391819 BX391819
     5     712.4     16.6    879  13  BU517328                  BU517328 AGENCOURT
     6     710.8     16.6    978  13  BX386457                  BX386457 BX386457
c    7     658.2     15.4    667  12  BQ028040                  BQ028040 UI-H-CO0-
c    8     649.6     15.2   1201  9   AL576908                  AL576908 AL576908
c    9     608.6     14.2    988  13  BX386456                  BX386456 BX386456
c   10     565.4     13.2    638  13  BU624356                  BU624356 UI-H-FG1-
    11       525     12.2    634  9   AA418644                  AA418644 zv93f07.r
    12     501.6     11.7    648  10  BB613975                  BB613975 BB613975
    13       498     11.6    613  14  CB158337                  CB158337 K-EST0217
    14       496     11.6    515  13  BX472602                  BX472602 DKFZp686C
    15     493.8     11.5    919  13  BX370169                  BX370169 BX370169
c   16     485.2     11.3    537  14  CA417651                  CA417651 UI-H-FE0-
    17     485.2     11.3    836  10  BG173799                  BG173799 602333853
c   18     474.6     11.1    544  10  BE143588                  BE143588 MR0-HT016
    19       473     11.0    473  13  BX279537                  BX279537 BX279537
    20       466     10.9    586  9   AA284110                  AA284110 zc36a07.T
    21     437.8     10.2   1495  11  AK042169                  AK042169 Mus muscu
    22     432.2     10.1    835  14  CD242054                  CD242054 AGENCOURT
c   23     424.8      9.9    456  9   AI278292                  AI278292 ql78d07.x
    24       414      9.7    437  9   AA132252                  AA132252 zo06h07.r
    25     408.6      9.5    480  12  BM537249                  BM537249 ha80a02.g
    26     400.4      9.3    490  12  BI975931                  BI975931 484508 MA
c   27     387.4      9.0    635  14  CA429478                  CA429478 UI-H-FH1-
c   28       387      9.0    722  14  CA420610                  CA420610 UI-H-FG0-
    29     384.4      9.0    431  10  BE006130                  BE006130 RC4-BN012
    30       383      8.9    853  10  BE573834                  BE573834 601331813
c   31     380.8      8.9    395  12  BG983552                  BG983552 IL5-CN006
    32       368      8.6    512  14  CA566520                  CA566520 K0402C12-
c   33     367.8      8.6    717  13  BQ775821                  BQ775821 UI-H-FH0-
c   34       356      8.3    687  13  BQ773484                  BQ773484 UI-H-FE0-
    35     354.4      8.3    476  13  BQ561040                  BQ561040 H4067H10-
    36     350.2      8.2    684  9   AV245523                  AV245523 AV245523
    37     345.2      8.1    806  12  BG919720                  BG919720 602822407
    38     338.8      7.9    511  14  CA893496                  CA893496 B0179D10-
c   39     338.4      7.9    652  12  BQ020033                  BQ020033 UI-H-ED0-
c   40     333.6      7.8    458  14  W46849                    W46849 zc36a07.s1
c   41     333.4      7.8    652  9   AI948463                  AI948463 wq06b02.x
    42       320      7.5    651  13  BX475936                  BX475936 DKFZp686G
    43     318.4      7.4    432  14  CB158287                  CB158287 K-EST0217
    44     312.4      7.3    324  14  L44418                    L44418 HUMEST1B5 H
    45     307.8      7.2    553  14  CA873206                  CA873206 K0923E11-

                                 ALIGNMENTS

RESULT 1
AK029247
LOCUS       AK029247                4485 bp    mRNA    linear   HTC 05-DEC-2002
DEFINITION  Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
            library, clone:4831440K17 product:weakly similar to
            alpha-2-macroglobulin (fragments) [Limulus polyphemus], full insert
            sequence.
ACCESSION   AK029247
VERSION     AK029247.1  GI:26081244
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
```

```
  PUBMED   11042159
REFERENCE  3
  AUTHORS  Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
           Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
           Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
           Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
           Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
           Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
           Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE    RIKEN integrated sequence analysis (RISA) system--384-format
           sequencing pipeline with 384 multicapillary sequencer
  JOURNAL  Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE  20530913
  PUBMED   11076861
REFERENCE  4
  AUTHORS  Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
           Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
           Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
           Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
           Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
           Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
           Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
           Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
           Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
           Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
           Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
           Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
           Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
           Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
           Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
           Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
           Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
           Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
           and Hayashizaki,Y.
  TITLE    Functional annotation of a full-length mouse cDNA collection
  JOURNAL  Nature 409 (6821), 685-690 (2001)
  MEDLINE  21085660
  PUBMED   11217851
REFERENCE  5
  AUTHORS  The FANTOM Consortium and the RIKEN Genome Exploration Research
           Group Phase I & II Team.
  TITLE    Analysis of the mouse transcriptome based on functional annotation
           of 60,770 full-length cDNAs
  JOURNAL  Nature 420, 563-573 (2002)
REFERENCE  6  (bases 1 to 4485)
  AUTHORS  Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
           Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
           Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
           Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
           Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
           Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
           Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
           Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
           Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
           Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
           Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
           Muramatsu,M. and Hayashizaki,Y.
  TITLE    Direct Submission
  JOURNAL  Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
           Physical and Chemical Research (RIKEN), Laboratory for Genome
           Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
           RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
           Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
           URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
           Fax:81-45-503-9216)
COMMENT    cDNA library was prepared and sequenced in Mouse Genome
           Encyclopedia Project of Genome Exploration Research Group in Riken
           Genomic Sciences Center and Genome Science Laboratory in RIKEN.
           Division of Experimental Animal Research in Riken contributed to
           prepare mouse tissues.
           Please visit our web site for further details.
           URL:http://genome.gsc.riken.go.jp/
           URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..4485
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:4831440K17"
                     /db_xref="taxon:10090"
                     /clone="4831440K17"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="0 day neonate"
     misc_feature    112..4439
                     /note="putative
                     weakly similar to alpha-2-macroglobulin (fragments)
                     [Limulus polyphemus] (PIR|A36260, evidence: FASTY,
                     51.8%ID, 76.7%length, match=168)"
BASE COUNT     1277 a    970 c   1010 g   1228 t
ORIGIN

  Query Match             63.2%;  Score 2708.8;  DB 11;  Length 4485;
  Best Local Similarity   77.9%;  Pred. No. 0;
  Matches 3393;  Conservative    0;  Mismatches  872;  Indels   91;  Gaps    7;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
Db        112 ATGCGCAGTCGGAGGCTCCTGAGCGCGGCCCACCTTCTCTGCTTGTGCGCGGTCGCGCTG 171

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
Db        172 GCC---GCCCCAGGGTCCCGCTTTCTGGTGACAGCCCCAGGGATCATCAGACCCGGAGCA 228

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
Db        229 AATGTGACTATTGGGGTGGATCTCCTGGAAAATAGCCCCCCACAGGTCCTTGTAAAGGCT 288

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
Db        289 CAGGTGTTCAAGATAGCTTCCAACAAATCGAGATCCATCCTGGAAGCAGAAGGAGTCTTT 348

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
Db        349 CACAGAGGCCATTTCAAGACTCTCGTTCTTCCGGCACTACCTCTGAGCAGTGCAGATAAG 408

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
Db        409 ATTTATGAGCTACACATAAACGGACAATCAGAGAATGAGATCGTATTCTCCAACAGGACA 468

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
Db        469 CGTTTAACATTTGAGAGCAAGAGTATATCTGTCCTCATTCAGACAGACAAGGCATTCTAC 528

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
Db        529 AAGCCGAAGCAGGAAGTGAAGTTTCGGGTGCTCACACTCTGCTCAGATCTCAAGCCTTAT 588

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
Db        589 AGGACATCCGTGGACATTTTCATTAAGGATCCGAAGTCTAATGTGATCCAACAGTGGTTT 648

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
Db        649 TCTCAGAAAGGTGATCTTGGAGTCGTTTCCAAAACTTTTCAACTATCCTCTAATCCAATA 708

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
Db        709 TTTGGTGACTGGTCCATTCAGGTTCAAGTGAATGATCAGCAATATTACCAGTCATTTCAG 768

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
Db        769 GTCCTAGAATATGTGTTACCAAAGTTCGAAGTCACCGTGCAGACCCCGCTGTATTGTTCC 828

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
```

```
Db    829 TTGAAGTCTAAACAGTTAAATGGCTCCGTCATAGCAAAGTATACATACGGGAAGCCAGTG 888

Qy    781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
Db    889 AAGGGAAGTCTGTCACTTACATTTTTACCTTTGTCTTTTTGGGGAAAGAAGAAGAATATC 948

Qy    841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
Db    949 ACAAAAAGTTTTGAGATAAATGGATTTGCAAACTTCTCTTTTGATAATTATGAGATGAAA 1008

Qy    901 AATGTAAT------------------GGATTCTTCAAATGGACTTTCTGAATACCTGGAT 942
Db   1009 AAGGTAATGAATTTGAAACCGCTAACGGATGTCTCGGAAGGGAGTTATGAGAATGTGGAC 1068

Qy    943 CTATCTTCCCCTGGACCAGTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATT 1002
Db   1069 CCGTCCTTTCCTGGGCCGGCAGAGATTATAGCTACTGTGACGGAATCACTCACAGGTATC 1128

Qy   1003 TCAAGAAATGTAAGCACTAATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTT 1062
Db   1129 TCAAGAATGGCAAGCACCAATGTGTTTTTCAAGCAACATGATTACATCATTGAGATTTTT 1188

Qy   1063 GATTATACTACTGTCTTGAAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGT 1122
Db   1189 GATTATACGACTGTCTTGAAACCATCTCTCAACTTCACAGCCACTGTGAAGGTCAGCCGC 1248

Qy   1123 GCTGATGGCAACCAACTGACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACA 1182
Db   1249 TCTGATGGCAATCAACTAACTCCTGAAGAAATAGAAAATGATTTGGTCACAGTAGTGACC 1308

Qy   1183 CAGAGAAACTATACTGAGTACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCT 1242
Db   1309 CAGGGAAAAATAACCACCCTGAGAGTCAGA----------GAGATCAGGAAATGGATTAT 1358

Qy   1243 GTTCAGAAAATAAATTATACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATC 1302
Db   1359 ATCCAGACCGTAAATTATACTATCCCCCAAAATGGAATCATTAAGATTGAATTCCCAGTC 1418

Qy   1303 CTGGAGGATTCCAGTGAGCTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATG 1362
Db   1419 ATGTCCATTTCCGGCGAGCTACAACTGAAGGCTTATTTTCTTGATGGTACAAGTTCTGTG 1478

Qy   1363 GCAGTTCATAGTCTGTTTAAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGAT 1422
Db   1479 ACAGTCCACAGTATGTTCACCTCTCCTAGTAAGACGTACATCCAGCTTAAAACCAGAGAT 1538

Qy   1423 GAAAATATAAAGGTGGGATCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAG 1482
Db   1539 GAATATATAAAGGTGGGGTCACCTTTTGATTTGATGGTTAGTGGCAACAGACAATTCAAG 1598

Qy   1483 GAGTTAAGCTATATGGTAGTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCA 1542
Db   1599 GACTTAAGCTATATGGTAATATCAAAGGGACAGTTGGTGGCTGCAGGCAAACAAAGTTCA 1658

Qy   1543 ACAATGTTCTCTTTAACACCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTAT 1602
Db   1659 AGAACTTTCTCTTTAACACCAGAGGCTTCCTGGGCTCCAAAGGCCTGTATAATCGCATAT 1718

Qy   1603 TATATTGAAGATGATGGGGAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTT 1662
Db   1719 TATATTGCAGAGGATGGGGAAATTATAAATGATATTCTAAAAATCCCCGTTCAGCTTGTT 1778

Qy   1663 TTTAAAAATAAGATAAAGCTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTC 1722
Db   1779 TTTGAAAATAAGGTAAAGCTGTTTTGGAGTAAACCTACAGTCAAGCCATCTGATAAGGTC 1838

Qy   1723 TCTCTTAGGATCTCTGTGACACAGCCTGACTCCCATAGTTGGGATTGTAGCTGTTGACAAA 1782
Db   1839 TCTCTCAGGATCTCTGCGACACAGTCTGACTCCCTGGTGGGGATTGTAGCTGTTGACAAA 1898

Qy   1783 AGTGTGAATCTGATGAATGCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTG 1842
Db   1899 AGTGTAACATTGATGGAAAACTCCAACAGTATTACAATGGAAACCATGGTCCATGAACTG 1958

Qy   1843 GAACTTTATAACACAGGATATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAG 1902
Db   1959 GAACTTTATAACACAGAATATTATCTAGGCATGTTCATGAACTCCTTCGCTGTATTTCAG 2018

Qy   1903 GAATGTGGACTCTGGGTATTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTT 1962
Db   2019 GAATGTGGTCTCTGGGTGTTGACAGATGCAACCCTTATAAGAGATAGCATTGATGAAGTC 2078

Qy   1963 TATGACAATGCAGAATATGCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGAT 2022
Db   2079 TACGATACTGAAGAGTATTCTGAACGCTTTGCAGAGGAAAATGAGGCAAACCTGGTTGAT 2138

Qy   2023 ATTCATGACTTTTCTTTGGGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGG 2082
Db   2139 TTTGAAGATGCTTCTTCAGTTAACAATGTACATGTCAGGAAGAATTTTCCAGAAACATGG 2198

Qy   2083 ATTTGGCTAGACACCAACATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCT 2142
Db   2199 ATTTGGCTAGACGCCTACATGGGATCTAAGATTTATGAAGAATTTGAAGTTACTGTACCT 2258

Qy   2143 GATTCTATCACTTCTTGGGTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGA 2202
Db   2259 GACTCCATCACTTCTTGGGTGGCTTCGGCTTTTGTCATCTCAGAAGACCTGGGTTTTGGG 2318

Qy   2203 CTAACAACTACTCCAGTGGAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTT 2262
Db   2319 TTAACAACCGTTCCAGCAGAGCTGCAAGCCTTCCAACCGTTTTTCCTTTTCCTGAATCTT 2378

Qy   2263 CCCTACTCTGTTATCAGAGGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTG 2322
Db   2379 CCGTACTCTGTTATCAGAGGTGAAGAGTTTGCTTTGGAAGTATCCATCGTCAATTATTTG 2438

Qy   2323 AAAGATGCCACTGAGGTTAAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATG 2382
Db   2439 AAAGATACAATTAAGGTTGTGATACTCATTGAGGAAAGTGACAGCTTTGATATTTTAATG 2498

Qy   2383 ACTTCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGAT 2442
Db   2499 ACTTCAAATGACACCAATGGCACCATATACCGGAAGACTGTTCAAGTTCCCAGGGATAAT 2558

Qy   2443 GGGGCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTC 2502
Db   2559 GGGGTCACTCTTGTTTTTCCCATCAAGCCCACGCACTTGGGAGAGATTCCCATCACAGTG 2618

Qy   2503 ACAGCTCTTTCACCCACTGCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAA 2562
Db   2619 ACGGCTGCCTCACCCACTGCCTCTGATGCTGTCACCCAGACAATTGTAGTAAAGCCTGAA 2678

Qy   2563 GGAATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAG 2622
Db   2679 GGGATAGAAAAGTCGTATTCGAAATCTGTCTTATTGGATCTGACAGATAGCAACGTAGAA 2738

Qy   2623 AGTACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGA 2682
Db   2739 AGTAAGCAGCAATCTATGAGATTCTCCTTTCCTCCTGATACGGTCATTGGCAGTGAAAGA 2798

Qy   2683 GTTCAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTG 2742
Db   2799 GTTCAGATCACAGCAATTGGAGATATCCTTGGTTCCTCCATCAATGGCTTATCTTCACTG 2858

Qy   2743 ATTCGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTAC 2802
Db   2859 ATCCGGATGCCTTATGGATGTGGTGAACAGAACATGATATATTTTGCTCCAAATATTTAC 2918

Qy   2803 ATTTTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTT 2862
Db   2919 ATTTTGGATTATCTGACTAAACAGAAACAGCTGACAGTTAATTTAAAAGAAAAGGCCCTT 2978

Qy   2863 TCATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTC 2922
Db   2979 TCATATATGAGGCAAGGTTACCAAAGGGAGCTTCTCTATCAGAGGGAAGATGGCTCCTTC 3038
```

```
Qy       2923 AGTGCTTTTGGGAATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGA 2982
Db       3039 AGTGCTTTTGGGGACATTGACTCTTCTGGGAGCACTTGGCTGTCAGCATTTGTTTTAAGA 3098

Qy       2983 TGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGAACATAC 3042
Db       3099 TGCTTTCTGGAAGCTGATTACTATATAGATATTGATCAGGATGTGTTACACAGAACATAT 3158

Qy       3043 ACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATT 3102
Db       3159 ACTTGGCTTAATGCACATAAGAAATTCAATGGTGAATTTTGGGAGCCAGGAAGAGTGATT 3218

Qy       3103 CATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACT 3162
Db       3219 CACAGTGAACTTCAAGGTGGCACCAAAAGCCCAGTAACACTTACGGCCTATATTGTGACT 3278

Qy       3163 TCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTT 3222
Db       3279 TCTGTCCTGGGATACAAAAAGTATCAGCCTAATATCGATGTACAAGACTCAATCAAGTTT 3338

Qy       3223 TTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTAT 3282
Db       3339 TTGGAATTTGAATTCAGCAGAGGAATTTCGGACAATTATACCTTAGCAATTATATCCTAT 3398

Qy       3283 GCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGA 3342
Db       3399 GCCCTGTCCACAGTCGGGAGCCCTAAAGCAGAGGAGGCTTTGAACTTGCTGATGCAGCGA 3458

Qy       3343 GCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCC 3402
Db       3459 TCAGAAAAGGAAGGAGACACACAGTTCTGGTTGTCATCAGGACCTGCGCTCTCTGGGTCC 3518

Qy       3403 TGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTA 3462
Db       3519 TGGCAGCCGCGCTCCGTGGACATCGAAATCGCTGCTTATGCACTGCTGGCGCACACTCTG 3578

Qy       3463 CAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGC 3522
Db       3579 CA---CCATGTGTCTGAGGGAATCCCGGTTATGAGGTGGCTCATCCAGCAAAGAAACAGC 3635

Qy       3523 TTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTT 3582
Db       3636 CTGGGAGGTTTTGTATCCACTCAGGATACTGTTGTGGCCCTGAAGGCATTATCTGAATTT 3695

Qy       3583 GCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCA 3642
Db       3696 TCAGCCCTTGTGCACAAGGAAAATACAGATATCCAACTGACCGTCACAGGGCCCGGCATC 3755

Qy       3643 CCAAGTCCT--------------------------------------------------- 3651
Db       3756 CCCAGATCCATACACTTCCGAATTGACTCCCAGAACCTCTTTCTCCTTCACCAGGAAGAG 3815

Qy       3652 CTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATT 3711
Db       3816 CTTCATGCGCTAGATCCCATCACCGTTAATGTTTCTGCGCATGGCTCGGGATTTGCTATT 3875

Qy       3712 TGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCT 3771
Db       3876 TGCCAGCTTAATGTTGACTATAACGTGAAAGGTTCAGGTTCTTCTAAAAGACGGAGATCT 3935

Qy       3772 ATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTC 3831
Db       3936 ACCGAAAATCAAGAAGTTTTTGATTTAGACGTCATTGTGA---ATAATGAGGACGACATT 3992

Qy       3832 AATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCT 3891
Db       3993 AGTCACCTGAATCTGAATGTGTGCACAAGTCACTTGGGTTCAGAGAGGACAGGCATGGTC 4052

Qy       3892 CTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGC 3951
Db       4053 CTTATGGAAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATTCAATTCCTCTGAGT 4112

Qy       3952 GAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTA 4011
Db       4113 GAGACCCTGAAGAAAGTGGAATATGATAACGGGAAACTCAACCTTTATTTAGATTCTGTA 4172

Qy       4012 AATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACC 4071
Db       4173 AATGAATCCCAGTTTTGTGTTAATATTCCCACTGTGAGAGACTACAAAGTTTCAAATATT 4232

Qy       4072 CAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGT 4131
Db       4233 CGAGATGGTTCAGTATCTGTAATGGACTACTATGAACCAAGGAGACAGGCAGTGCGAAGC 4292

Qy       4132 TACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGT 4191
Db       4293 TACAACACCCAGGTGAAGCTGTCCTCATGTTACCT---CAGTCCAGACACCAACTGCAAG 4349

Qy       4192 CCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGT 4251
Db       4350 TCACACACGGACGGAGCCACGGACTCCCTTCGACGTTCTTCGAGCCTTCTCGTCTTTTGC 4409

Qy       4252 TTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 4287
Db       4410 TCCGTCCTTCTGTACTTTGTGCAACATTGATTATGA 4445
```

RESULT 2
AK036799

```
LOCUS       AK036799                4725 bp    mRNA    linear   HTC 05-DEC-2002
DEFINITION  Mus musculus adult female vagina cDNA, RIKEN full-length enriched
            library, clone:9930012E15 product:weakly similar to
            alpha-2-macroglobulin (fragments) [Limulus polyphemus], full insert
            sequence.
ACCESSION   AK036799
VERSION     AK036799.1  GI:26085434
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
```

```
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
   MEDLINE  21085660
   PUBMED   11217851
REFERENCE   5
  AUTHORS   The FANTOM Consortium and the RIKEN Genome Exploration Research
            Group Phase I & II Team.
  TITLE     Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL   Nature 420, 563-573 (2002)
REFERENCE   6  (bases 1 to 4725)
  AUTHORS   Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
            Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
            Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
            Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
            Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
            Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
            Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
            Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
            Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
            Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.go.jp/
            URL:http://fantom.gsc.riken.go.jp/.
FEATURES             Location/Qualifiers
     source          1..4725
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9930012E15"
                     /db_xref="taxon:10090"
                     /clone="9930012E15"
                     /sex="female"
                     /tissue_type="vagina"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     misc_feature    97..4429
                     /note="putative
                     weakly similar to alpha-2-macroglobulin (fragments)
                     [Limulus polyphemus] (PIR|A36260, evidence: FASTY,
                     51.8%ID, 76.7%length, match=168)"
     polyA_signal    4707..4712
                     /note="putative"
     polyA_site      4725
                     /note="putative"
BASE COUNT     1334 a   1016 c   1066 g   1309 t
```

```
ORIGIN

  Query Match             62.3%;  Score 2668.8;  DB 11;  Length 4725;
  Best Local Similarity   77.7%;  Pred. No. 0;
  Matches 3389;  Conservative    0;  Mismatches  877;  Indels   94;  Gaps   10;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||   | |    ||||||| ||| ||||||| |||||| ||||| | | |||||||
Db         97 ATGCGCAGTCGGAGGCTCCTGAGCGCGGCCCACCTTCTCTGCTTGTGCGCGGTCGCGCTG 156

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              |||    || || |||  | || |||||||||||||||||||||||||||| ||||||| |
Db        157 GCC---GCCCCAGGGTCCCGCTTTCTGGTGACAGCCCCAGGGATCATCAGACCCGGAGCA 213

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||| || |||||| |     ||||  ||||||    ||| |||||
Db        214 AATGTGACTATTGGGGTGGATCTCCTGGAAAATAGCCCCCCACAGGTCCTTGTAAAGGCT 273

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
               || || |||||| ||| || |||     |      ||  ||||||||||||||||||||
Db        274 CAGGTGTTCAAGATAGCTTCCAACAAATCGAGATCCATCCTGGAAGCAGAAGGAGTCTTT 333

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
                | | ||||  ||| ||||| ||   ||||||  ||||||||||| ||||||||| ||
Db        334 CACAGAGGCCATTTCAAGACTCTCGTTCTTCCGGCACTACCTCTGAGCAGTGCAGATAAG 393

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||    ||| |||||   |  || ||||||| ||||||| || || ||
Db        394 ATTTATGAGCTACACATAAACGGACAATCAGAGAATGAGATCGTATTCTCCAACAGGACA 453

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              || ||| |||||||||  ||||| ||||||||| |||||| ||||||||||||| || |||
Db        454 CGTTTAACATTTGAGAGCAAGAGTATATCTGTCCTCATTCAGACAGACAAGGCATTCTAC 513

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||| ||||| ||||||||||||||    |  ||||||| |||||| | |||||||
Db        514 AAGCCGAAGCAGGAAGTGAAGTTTCGGGTGCTCACACTCTGCTCAGATCTCAAGCCTTAT 573

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              |  || ||   |  ||||| |||||||||| || || || ||| |||||||||||||||
Db        574 AGGACATCCGTGGACATTTTCATTAAGGATCCGAAGTCTAATGTGATCCAACAGTGGTTT 633

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              || ||  || ||||||||||||| |||||||||||||||| |||||  ||  ||||||||
Db        634 TCTCAGAAAGGTGATCTTGGAGTCGTTTCCAAAACTTTTCAACTATCCTCTAATCCAATA 693

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
               ||||||||||||| |||| ||| ||||||||||||| |||  |||||| ||  ||||||||
Db        694 TTTGGTGACTGGTCCATTCAGGTTCAAGTGAATGATCAGCAATATTACCAGTCATTTCAG 753

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||   |||||||| |||||||| || ||||| || ||||||| ||   | ||||||||
Db        754 GTCCTAGAATATGTGTTACCAAAGTTCGAAGTCACCGTGCAGACCCCGCTGTATTGTTCC 813

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
               |||| ||||| || ||||||||  || |||  |||||||||||||| |||||| |
Db        814 TTGAAGTCTAAACAGTTAAATGGCTCCGTCATAGCAAAGTATACATACGGGAAGGCCAGT 873

Qy        781 AAAGG--AGACGTAACGCTTACATTTTTACCTTTATCCTTTT-GGGGAAAGAAGAAAAAT 837
               |||| ||    |  | ||||||||||||||| || |||| ||||||||||||| |||
Db        874 GAAGGGAAGGTCTGTCACTTACATTTTTACCTTTGTCTTTTTGGGGGAAAGAAGAAGAAT 933

Qy        838 ATTACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATG 897
              || |||||||  ||| ||||||||||||| |||||||||||||||| || || | |||||
Db        934 ATCACAAAAAGTTTTGAGATAAATGGATTTGCAAACTTCTCTTTTGATAATTATGAGATG 993

Qy        898 AAAAATGTAAT-----------------GGATTCTTCAAATGGACTTTCTGAATACCTG 939
              ||||| |||||                 ||||    ||  | ||    || |||  ||
Db        994 AAAAAGGTAATGAATTTGAAACCGCTAACGGATGTCTCGGAAGGGAGTTATGAGAATGTG 1053

Qy        940 GATCTATCTTCCCCTGGACCAGTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGT 999
              || |  || |  ||||| || | ||| ||| || | || ||||| |||||| | ||||||
Db       1054 GACCCGTCCTTTCCTGGGCCGGCAGAGATTATAGCTACTGTGACGGAATCACTCACAGGT 1113
```

```
Qy   1000 ATTTCAAGAAATGTAAGCACTAATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTT 1059
Db   1114 ATCTCAAGAATGGCAAGCACCAATGTGTTTTTCAAGCAACATGATTACATCATTGAGATT 1173

Qy   1060 TTTGATTATACTACTGTCTTGAAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACT 1119
Db   1174 TTTGATTATACGACTGTCTTGAAACCATCTCTCAACTTCACAGCCACTGTGAAGGTCAGC 1233

Qy   1120 CGTGCTGATGGCAACCAACTGACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAG-T 1178
Db   1234 CGCTCTGATGGCAATCAACTAACTCCTGAAGAAATAGAAAATGATTTGGTCACAGTAGCT 1293

Qy   1179 GACACAGAGAAACTATACTGAGTACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGA 1238
Db   1294 GACCCAGCGGAAAAAATAACCACCCTGAGAG---------TCAGAGAGATCAGGAAATGGA 1344

Qy   1239 AGCTGTTCAGAAAATAAATTATACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCC 1298
Db   1345 TTATATCCAGACCGTAAATTATACTATCCCCCAAAATGGAATCATTAAGATTGAATTCCC 1404

Qy   1299 AATCCTGGAGGATTCCAGTGAGCTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAG 1358
Db   1405 AGTCATGTCCATTTCCGGCGAGCTACAACTGAAGGCTTATTTTCTTGATGGTACAAGTTC 1464

Qy   1359 CATGGCAGTTCATAGTCTGTTTAAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAG 1418
Db   1465 TGTGACAGTCCACAGTATGTTCACCTCTCCTAGTAAGACGTACATCCAGCTTAAAACCAG 1524

Qy   1419 AGATGAAAATATAAAGGTGGGATCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATT 1478
Db   1525 AGATGAATATATAAAGGTGGGGTCACCTTTTGATTTGATGGTTAGTGGCAACAGACAATT 1584

Qy   1479 GAAGGAGTTAAGCTATATGGTAGTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAA 1538
Db   1585 CAAGGACTTAAGCTATATGGTAATATCAAAGGGACAGTTGGTGGCTGCAGGCAAACAAAG 1644

Qy   1539 TTCAACAATGTTCTCTTTAACACCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGT 1598
Db   1645 TTCAAGAACTTTCTCTTTAACACCAGAGGCTTCCTGGGCTCCAAAGGCCTGTATAATCGC 1704

Qy   1599 GTATTATATTGAAGATGATGGGGAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCT 1658
Db   1705 ATATTATATTGCAGAGGATGGGGAAATTATAAATGATATTCTAAAAATCCCCGTTCAGCT 1764

Qy   1659 TGTTTTTAAAAATAAGATAAAGCTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAA 1718
Db   1765 TGTTTTTGAAAATAAGGTAAAGCTGTTTTGGAGTAAACCTACAGTCAAGCCATCTGATAA 1824

Qy   1719 AGTCTCTCTTAGGATCTCTGTGACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGA 1778
Db   1825 GGTCTCTCTCAGGATCTCTGCGACACAGTCTGACTCCCTGGTGGGGATTGTAGCTGTTGA 1884

Qy   1779 CAAAAGTGTGAATCTGATGAATGCCTCTAATGATATTACAATGGAAAATGTGGTCCATGA 1838
Db   1885 CAAAAGTGTAACATTGATGGAAAACTCCAACAGTATTACAATGGAAACCATGGTCCATGA 1944

Qy   1839 GTTGGAACTTTATAACACAGGATATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTT 1898
Db   1945 ACTGGAACTTTATAACACAGAATATTATCTAGGCATGTTCATGAACTCCTTCGCTGTATT 2004

Qy   1899 TCAGGAATGTGGACTCTGGGTATTGACAGATGCAAACCTCACGAAGGATTATATTGATGG 1958
Db   2005 TCAGGAATGTGGTCTCTGGGTGTTGACAGATGCAACCCTTATAAGAGATAGCATTGATGA 2064

Qy   1959 TGTTTATGACAATGCAGAATATGCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGT 2018
Db   2065 AGTCTACGATACTGAAGAGTATTCTGAACGCTTTGCAGAGGAAAATGAGGCAAACCTGGT 2124

Qy   2019 AGATATTCATGACTTTTCTTTGGGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGAC 2078
Db   2125 TGATTTTGAAGATGCTTCTTCAGTTAACAATGTACATGTCAGGAAGAATTTTCCAGAAAC 2184

Qy   2079 TTGGATTTGGCTAGACACCAACATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGT 2138
Db   2185 ATGGATTTGGCTAGACGCCTACATGGGATCTAAGATTTATGAAGAATTTGAAGTTACTGT 2244

Qy   2139 ACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCT 2198
Db   2245 ACCTGACTCCATCACTTCTTGGGTGGCTTCGGCTTTTGTCATCTCAGAAGACCTGGGTTT 2304

Qy   2199 TGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAA 2258
Db   2305 TGGGTTAACAACCGTTCCAGCAGAGCTGCAAGCCTTCCAACCGTTTTTCCTTTTCCTGAA 2364

Qy   2259 TCTTCCCTACTCTGTTATCAGAGGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTA 2318
Db   2365 TCTTCCGTACTCTGTTATCAGAGGTGAAGAGTTTGCTTTGGAAGTATCCATCGTCAATTA 2424

Qy   2319 TTTGAAAGATGCCACTGAGGTTAAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCT 2378
Db   2425 TTTGAAAGATACAATTAAGGTTGTGATACTCATTGAGGAAAGTGACAGCTTTGATATTTT 2484

Qy   2379 AATGACTTCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGA 2438
Db   2485 AATGACTTCAAATGACACCAATGGCACCATATACCGGAAGACTGTTCAAGTTCCCAGGGA 2544

Qy   2439 GGATGGGGCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCTATCAC 2498
Db   2545 TAATGGGGTCACTCTTGTTTTTCCCATCAAGCCCACGCACTTGGGAGAGATTCCCATCAC 2604

Qy   2499 AGTCACAGCTCTTTCACCCACTGCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGC 2558
Db   2605 AGTGACGGCTGCCTCACCCACTGCCTCTGATGCTGTCACCCAGACAATTGTAGTAAAGCC 2664

Qy   2559 TGAAGGAATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAATAGGCT 2618
Db   2665 TGAAGGGATAGAAAAGTCGTATTCGAAATCTGTCTTATTGGATCTGACAGATAGCAACGT 2724

Qy   2619 ACAGAGTACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGA 2678
Db   2725 AGAAAGTAAGCAGCAATCTATGAGATTCTCCTTTCCTCCTGATACGGTCATTGGCAGTGA 2784

Qy   2679 AAGAGTTCAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTC 2738
Db   2785 AAGAGTTCAGATCACAGCAATTGGAGATATCCTTGGTTCCTCCATCAATGGCTTATCTTC 2844

Qy   2739 ATTGATTCGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCAAATAT 2798
Db   2845 ACTGATCCGGATGCCTTATGGATGTGGTGAACAGAACATGATATATTTTGCTCCAAATAT 2904

Qy   2799 TTACATTTTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGC 2858
Db   2905 TTACATTTTGGATTATCTGACTAAACAGAAACAGCTGACAGTTAATTTAAAAGAAAAGGC 2964

Qy   2859 TCTTTCATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTC 2918
Db   2965 CCTTTCATATATGAGGCAAGGTTACCAAAGGGAGCTTCTCTATCAGAGGGAAGATGGCTC 3024

Qy   2919 TTTCAGTGCTTTTGGGAATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTT 2978
Db   3025 CTTCAGTGCTTTTGGGGACATTGACTCTTCTGGGAGCACTTGGCTGTCAGCATTTGTTTT 3084

Qy   2979 AAGATGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGAAC 3038
Db   3085 AAGATGCTTTCTGGAAGCTGATTACTATATAGATATTGATCAGGATGTGTTACACAGAAC 3144

Qy   3039 ATACACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGT 3098
Db   3145 ATATACTTGGCTTAATGCACATAAGAAATTCAATGGTGAATTTTGGGAGCCAGGAAGAGT 3204

Qy   3099 GATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGT 3158
Db   3205 GATTCACAGTGAACTTCAAGGTGGCACCAAAAGCCCAGTAACACTTACGGCCTATATTGT 3264

Qy   3159 AACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCA 3218
```

```
Db     3265 GACTTCTGTCCTGGGATACAAAAAGTATCAGCCTAATATCGATGTACAAGACTCAATCAA 3324

Qy     3219 TTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCTAGCCCTTATAAC 3278
Db     3325 GTTTTGGAATTTGAATTCAGCAGAGGAATTTCGGACAATTATACCTTAGCAATTATATC 3384

Qy     3279 TTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTG 3338
Db     3385 CTATGCCCTGTCCACAGTCGGGAGCCCTAAAGCAGAGGAGGCTTTGAACTTGCTGATGCA 3444

Qy     3339 GAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGA 3398
Db     3445 GCGATCAGAAAAGGAAGGAGACACACAGTTCTGGTTGTCATCAGGACCTGCGCTCTCTGG 3504

Qy     3399 CTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTT 3458
Db     3505 GTCCTGGCAGCCGCGCTCCGTGGACATCGAAATCGCTGCTTATGCACTGCTGGCGCACAC 3564

Qy     3459 CTTACAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAA 3518
Db     3565 TCTGCA---CCATGTGTCTGAGGGAATCCCGGTTATGAGGTGGCTCATCCAGCAAAGAAA 3621

Qy     3519 TAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGA 3578
Db     3622 CAGCCTGGGAGGTTTTGTATCCACTCAGGATACTGTTGTGGCCCTGAAGGCATTATCTGA 3681

Qy     3579 ATTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAG 3638
Db     3682 ATTTTCAGCCCTTGTGCACAAGGAAAATACAGATATCCAACTGACCGTCACAGGGCCCGG 3741

Qy     3639 CTCACCAAGTCCT------------------------------------------------ 3651
Db     3742 CATCCCCAGATCCATACACTTCCGAATTGACTCCCAGAACCTCTTTCTCCTTCACCAGGA 3801

Qy     3652 ----CTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGC 3707
Db     3802 AGAGCTTCATGCGCTAGATCCCATCACCGTTAATGTTTCTGCGCATGGCTCGGGATTTGC 3861

Qy     3708 TATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAG 3767
Db     3862 TATTTGCCAGCTTAATGTTGACTATAACGTGAAAGGTTCAGGTTCTTCTAAAAGACGGAG 3921

Qy     3768 ATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGA 3827
Db     3922 ATCTACCGAAAATCAAGAAGTTTTTGATTTAGACGTCATTGTGA---ATAATGAGGACGA 3978

Qy     3828 TCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCAT 3887
Db     3979 CATTAGTCACCTGAATCTGAATGTGTGCACAAGTCACTTGGGTTCAGAGAGGACAGGCAT 4038

Qy     3888 GGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCT 3947
Db     4039 GGTCCTTATGGAAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATTCAATTCCTCT 4098

Qy     3948 GAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTC 4007
Db     4099 GAGTGAGACCCTGAAGAAAGTGGAATATGATAACGGGAAACTCAACCTTTATTTAGATTC 4158

Qy     4008 TGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAA 4067
Db     4159 TGTAAATGAATCCCAGTTTTGTGTTAATATTCCCACTGTGAGAGACTACAAAGTTTCAAA 4218

Qy     4068 TACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAG 4127
Db     4219 TATTCGAGATGGTTCAGTATCTGTAATGGACTACTATGAACCAAGGAGACAGGCAGTGCG 4278

Qy     4128 AAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTG 4187
Db     4279 AAGCTACAACACCCAGGTGAAGCTGTCCTCGTGTTACCT---CAGTCCAGACACCAACTG 4335

Qy     4188 CCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTT 4247
Db     4336 CAAGTCACACACGGACGGAGCCACGGACTCCCTTCGACGTTCTTCGAGCCTTCTCGTCTT 4395

Qy     4248 CTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 4287
Db     4396 TTGCTCCGTCCTTCTGTACTTTGTGCAACATTGATTATGA 4435
```

```
RESULT 3
AL554092
LOCUS       AL554092                1201 bp    mRNA    linear   EST 31-MAY-2003
DEFINITION  AL554092 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 5-PRIME, mRNA sequence.
ACCESSION   AL554092
VERSION     AL554092.2  GI:31275905
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 15, 2001 this sequence version replaced gi:12894544.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI081DA09QP1&cluster=3578.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI081DA09QP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      359 a    238 c    236 g    293 t     75 others
ORIGIN

  Query Match             20.8%;  Score 893;  DB 9;  Length 1201;
  Best Local Similarity   93.9%;  Pred. No. 2.3e-187;
  Matches  947;  Conservative   39;  Mismatches   14;  Indels    9;  Gaps    6;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
Db         92 ATGCAGGGCCCACCGCTCCTGAMCGCMGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 151

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
Db        152 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGAMAGCMCCAGGGATMATCAGGCCCGGAGGA 211

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
Db        212 AATGTGACTATTGGGGTGGAGCTTCTGGAAMACTGCCCTTMAMAGGTGACTGTGAAGGCG 271

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
Db        272 GAGCTGCTCAAGACAGCATCAAAMMTCACTGTCTCTGTMCTGGAAGCAGAAGGAGTCTTT 331

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
```

```
Db        332 GAAAAARGCTCTTTTAARAMACTTAMTCTTCCATMAMTACMTCTGAACARTGCAGATGAG 391

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||:|||||||||||||||||||||||||||||||||||||||||:|||||
Db        392 ATTTATGAGCTAMGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAAWAGTACC 451

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||:||||||||||||||:||||||||||||
Db        452 CGCTTATCATTTGAGACCAAGAGAATATCTGTMTTCATTCAAACAGAMAAGGCCTTATAC 511

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        512 AAGCAAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 571

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        572 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 631

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        632 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 691

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        692 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 751

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        752 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 811

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db        812 ATGAATTCTAAGCATTTAAATGGTA-CATCACGGCAAAGTATACATATGGGAAGCCAGTG 870

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        871 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 930

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||:||||:||||||||||||
Db        931 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTWATGAWGAAGAGATGAAA 990

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              |:  :|   |  |||||||||||||||||||| ||::||||:|||:|  |||  |  |:
Db        991 AWGTWA--TGGATCTTCAAATGGACTTTCTGAA-ACYKGGATYTATYT--CCCCTGGACM 1045

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAA 1009
              ||||||||||| ||||||||||:|||  ::: ||:|||:|||  || |:
Db       1046 GTAGAAATTTT-ACCACAGTGAMAGA--YMKGTAMAGGWATTMAAAAAW 1091
```

RESULT 4
BX391819/c

```
LOCUS       BX391819                 941 bp    mRNA    linear   EST 13-MAY-2003
DEFINITION  BX391819 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 3-PRIME, mRNA sequence.
ACCESSION   BX391819
VERSION     BX391819.1  GI:30627404
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 941)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAI029ZD07_CS02740_1&cluster=3578.f.
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0BAI029ZD07_CS02740_1.
FEATURES             Location/Qualifiers
     source          1..941
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      271 a    214 c     181 g     266 t       9 others
ORIGIN

  Query Match             20.4%;  Score 872.8;  DB 13;  Length 941;
  Best Local Similarity   96.7%;  Pred. No. 6.8e-183;
  Matches  909;  Conservative    0;  Mismatches   28;  Indels    3;  Gaps    2;

Qy       2820 TAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGG 2879
              |||||  || || |||||||||| ||  ||||||||      ||||||     |||||||
Db        939 TAAAAGNAANCANCTGACAGATATTTGAAAAGAAAAGCTCTNTTCATTATGAGGGCAAGG 880

Qy       2880 TTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTA 2939
               |||||||||||  ||||||||||||||||||||||||| ||||||||||||||||||||
Db        879 GTACCAGAGAGACTTTCTCTATCAGAGGGAAGATGGCTC-TTCAGTGCTTTTGGGAATTA 821

Qy       2940 TGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGA 2999
              |||||||||||||||||||||||||||||||| |||| ||||||||||||||||||||||
Db        820 TGACCCTTCTGGGAGCACTTGGTTGTCAGCTTNTGTTNTAAGATGTTTCCTTGAAGCCGA 761

Qy       3000 TCCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACA 3059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        760 TCCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACA 701

Qy       3060 TCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGG 3119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        700 TCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGG 641

Qy       3120 TGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAG 3179
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 TGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAG 581

Qy       3180 AAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAG 3239
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        580 AAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAG 521

Qy       3240 TAGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGG 3299
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        520 TAGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGG 461

Qy       3300 GAGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGG 3359
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 GAGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGG 401

Qy       3360 CATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCT 3419
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 CATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCT 341

Qy       3420 GGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGA 3479
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 GGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGA 281

Qy       3480 GGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATC 3539
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 GGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATC 221
```

```
Qy       3540 TACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATAC 3599
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 TACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATAC 161

Qy       3600 AGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGT 3659
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 AGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGT 101

Qy       3660 GGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCT 3719
              ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 GGTACAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCT 41

Qy       3720 CAATG--TTGTATATAATGTGAAGGCTTCTGGGTCTTCTA 3757
              |||||  ||||||||||||| | |  ||||||||||||||
Db         40 CAATGNNTTGTATATAATGTAAGGCGTTCTGGGTCTTCTA 1


RESULT 5
BU517328
LOCUS       BU517328                 879 bp    mRNA    linear   EST 12-SEP-2002
DEFINITION  AGENCOURT_10164521 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6514905
            5', mRNA sequence.
ACCESSION   BU517328
VERSION     BU517328.1  GI:22824854
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM14090  row: p  column: 10
            High quality sequence stop: 663.
FEATURES             Location/Qualifiers
     source          1..879
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6514905"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_71"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2.1 kb.  "
BASE COUNT      248 a    171 c    206 g    253 t     1 others
ORIGIN

  Query Match             16.6%;  Score 712.4;  DB 13;  Length 879;
  Best Local Similarity   91.8%;  Pred. No. 3e-147;
  Matches  807;  Conservative    0;  Mismatches   17;  Indels   55;  Gaps    3;

Qy       3207 AGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCT 3266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCT 60

Qy       3267 AGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAA 3326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAA 120

Qy       3327 TATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTC 3386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTC 180

Qy       3387 CAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACT 3446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACT 240

Qy       3447 GCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAG 3506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAG 300

Qy       3507 CAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAA 3566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAA 360

Qy       3567 GGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGT 3626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGT 420

Qy       3627 GACGGGGCCTAGCTCACCAAGTCCT----------------------------------- 3651
              |||||||||||||||||||||||||
Db        421 GACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTTCTGATTGACACACACAACCGCTTACT 480

Qy       3652 ----------------CTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGG 3695
                              ||||||||||||||||||||||||||||||||||||||||||||
Db        481 CCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGG 540

Qy       3696 TTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTC 3755
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTC 600

Qy       3756 TAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGA 3815
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGA 660

Qy       3816 AAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGG 3875
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGG 720

Qy       3876 TAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGA 3935
              ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db        721 TAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGGCTTATGGTGCCTTCAGA 780

Qy       3936 AGCAATTTCT-CTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACC 3994
              |||||||||| |||||||||||||||| |||||||| ||||||||||| |||     | |
Db        781 AGCAATTTCTCCTGAGCGAGACAGTGAGGAAAGTGGGATATGATCATGGGAAAACTCAAC 840

Qy       3995 TCT---ATTTAGATTCTGTAAATGAAACCCAGTTTTGTG 4030
               ||    || |||||||||||||||||||    ||| ||
Db        841 CCTCTATTTNAGATTCTGTAAATGAAACCCCAGTTTTTG 879


RESULT 6
BX386457
LOCUS       BX386457                 978 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  BX386457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 5-PRIME, mRNA sequence.
ACCESSION   BX386457
VERSION     BX386457.1  GI:30457430
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 978)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
```

```
                 more information about this cluster, see
                 http://www.genoscope.cns.fr/
                 cgi-bin/cluster.cgi?seq=CS1AI021ZB10QP1&cluster=3578.f. Contact :
                 Feng Liang Email : fliang@lifetech.com URL :
                 http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
                 Faraday Avenue Genoscope sequence ID : CS1AI021ZB10QP1.
FEATURES             Location/Qualifiers
     source          1..978
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      374 a    151 c    179 g    220 t     54 others
ORIGIN

  Query Match             16.6%;  Score 710.8;  DB 13;  Length 978;
  Best Local Similarity   84.6%;  Pred. No. 6.9e-147;
  Matches  758;  Conservative   46;  Mismatches   90;  Indels    2;  Gaps    2;

Qy          7 GGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTGGCCGTG 66
              || ||| ||||||||| ||  || ||  ||||||||||||||||||||||||||||||||
Db         84 GGGCCANCGCTCCTGAACGNAGCACAAATCCTCTGCGTGTGCACCGCCGCGCTGGCCGTG 143

Qy         67 GCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGAAATGTG 126
              |||||:||||:|||||||:||||||:|||:|||||||| ||:|||| |||||||||||||
Db        144 GCTCCMGGGCMTCGGTTTMTGGTGAMAGCMCCAGGGATAATMAGGCACGGAGGAAATGTG 203

Qy        127 ACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCGGAGCTG 186
              ||:|||||||||||||||||||||||||||:||||||||||||||||||||||||| ::|
Db        204 ACWATTGGGGTGGAGCTTCTGGAACACTGCMCTTCACAGGTGACTGTGAAGGCGGA-SWG 262

Qy        187 CTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTTGAAAAA 246
              ||||||||||||| ||| :|:|  || || ||  |||||| |||| ||||:|||||||||
Db        263 CTCAAGACAGCATAAAAAMTMAAAGTATCAGTAATGGAAGAAGAAAGAGTMTTTGAAAAA 322

Qy        247 GGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAGATTTAT 306
               | |||:||||:| | : | :  || |  |  |   : :  |||||||||||||:|:| ||
Db        323 AGATCTWTTAARAAAAWAAAWAATCAAAAAAAAAAAWAWAAACAGTGCAGATGARAWTAAT 382

Qy        307 GAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACCCGCTTA 366
               |:|:| ||||||  ||| |||||||||||||:|:|:|| |||||||:|| || :||:||
Db        383 AARCWAAGTGTAAAAGGAAGTACCCAGGATGARAWTWTAATCTCTAAWAGAACAMGCWTA 442

Qy        367 TCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATACAAGCCA 426
                | |||| ||||| | ||||      |   :| |    ||| |:| ||||   ||| ||||||||
Db        443 AAAATTGAAACCAAAAAAATAAAAGAAWTAAAAAAAAAAARAAAAGGAATTAAACAAGCCA 502

Qy        427 AAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTACAAAACC 486
              ||||||||||||||||||||||||||||||||||||| |:| || |||| ||||||||  
Db        503 AAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTAARAATTAAAGCATTACAAAAAA 562

Qy        487 TCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTGTCACAA 546
               ||||||||||| ||||||||||  |||||| ||||||||||||||||||||||||||||
Db        563 ACTTTAAACATTATCATTAAGGAAACCAAATAAAATTTGATCCAACAGTGGTTGTCACAA 622

Qy        547 CAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATACTTGGT 606
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        623 CAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATACTTGGT 682

Qy        607 GACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAGGTTTCA 666
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        683 GACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAGGTTTCA 742

Qy        667 GAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCTATGAAT 726
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        743 GAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCTATGAAT 802

Qy        727 TCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTGAAAGGA 786
              ||||||||||||||||||||:||||  |||||: ||| | || : ||| :||:||||  |
Db        803 TCTAAGCATTTAAATGGTACYATCATTGCAAARSATATAGATTSTAAGTYAGYGAAATCA 862

Qy        787 GACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGG-AAAGAAGAAAAATATTACAAA 845
              ||| |||| |||:|||||||:||||||::||:| ||||| |||:||:||||||||||||||
Db        863 GACTTAACTCTKACATTTTKACCTTTWKCCKTGTGGGGSAAAKAAKAAAAATATTACAAA 922

Qy        846 AACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAA 901
              |||||||||||||||||:|||||||||:||||||||||||||||||||||||||||
Db        923 AACATTTAAGATAAATGSATCTGCAAAYTTCTCTTTTAATGATGAAGAGATGAAAA 978


RESULT 7
BQ028040/c
LOCUS       BQ028040                 667 bp    mRNA    linear   EST 27-MAR-2002
DEFINITION  UI-H-CO0-arh-c-08-0-UI.s1 NCI_CGAP_Sub9 Homo sapiens cDNA clone
            IMAGE:3106526 3', mRNA sequence.
ACCESSION   BQ028040
VERSION     BQ028040.1  GI:19763319
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 667)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jose Mercuende
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1..667
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3106526"
                     /tissue_type="mixed"
                     /dev_stage="mixed"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_Sub9"
                     /note="Vector: pT7T3-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoR I; Site_2: Not I; tissues:
                     Cholonic mucosa with Crohns disease, Cholonic mucosa with
                     ulcerative colitis, Fetal thymus, Cervix, Cervical
                     adenosquamous carcinoma, Ligament cells, Prostate
                     carcinoma, Bladder carcinoma, Brain oligodenroga ;
                     NCI_CGAP_Sub9 is a subtracted cDNA library constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tags for this library are CGTC, AACG, GGGCC,
                     GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional
                     information, contact: Bento Soares, bento-soares@uiowa.edu
                     TAG_LIB=UI-H-CO0
                     TAG_TISSUE=Cervical Adenosquamous Carcinoma
                     TAG_SEQ=CGAAG"
BASE COUNT      194 a    113 c    139 g    221 t
ORIGIN
```

```
  Query Match             15.4%;  Score 658.2;  DB 12;  Length 667;
  Best Local Similarity   99.5%;  Pred. No. 3.1e-135;
  Matches  660;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        176 AGGCGGAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAG 235
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 AGGCGGAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAG 608

Qy        236 TCTTTGAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAG 295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 TCTTTGAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAG 548

Qy        296 ATGAGATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATA 355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 ATGAGATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATA 488

Qy        356 GTACCCGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCT 415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 GTACCCGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCT 428

Qy        416 TATACAAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGC 475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 TATACAAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGC 368

Qy        476 CTTACAAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGT 535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 CTTACAAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGT 308

Qy        536 GGTTGTCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATC 595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 GGTTGTCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATC 248

Qy        596 CAATACTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCAT 655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 CAATACTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCAT 188

Qy        656 TTCAGGTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATT 715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 TTCAGGTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATT 128

Qy        716 GTTCTATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGC 775
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 GTTCTATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGC 68

Qy        776 CAGTGAAAGGAGACGTAACGCTTACATTTTACCTTTATCCTTTTGGGGAAAGAAGAAAA 835
              |||||||||||||||||||||||||||||||||||||||||||||||||||| || ||||
Db         67 CAGTGAAAGGAGACGTAACGCTTACATTTTACCTTTATCCTTTTGGGGAAAAAAAAAAA 8

Qy        836 ATA 838
              | |
Db          7 AAA 5


RESULT 8
AL576908/c
LOCUS       AL576908                1201 bp    mRNA    linear   EST 01-JUN-2003
DEFINITION  AL576908 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 3-PRIME, mRNA sequence.
ACCESSION   AL576908
VERSION     AL576908.2  GI:31315187
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1201)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     On Feb 16, 2001 this sequence version replaced gi:12939516.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0DI081DA09NP1&cluster=3578.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0DI081DA09NP1.
FEATURES             Location/Qualifiers
     source          1. .1201
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      344 a    248 c    224 g    321 t     64 others
ORIGIN

  Query Match             15.2%;  Score 649.6;  DB 9;  Length 1201;
  Best Local Similarity   94.7%;  Pred. No. 2.9e-133;
  Matches  679;  Conservative   11;  Mismatches   25;  Indels    2;  Gaps    2;

Qy       3571 CTGTCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTGACG 3630
              :| | : :|  ||||||||||||||||:|||| ||:|||||||||||||||||||||| 
Db        996 YTCTTYTRAATTGCAGCCCTAATGAAWACAG-AARGACAAATATCCAAGTGACCGTGAC- 939

Qy       3631 GGCCTAGCTCACCAAGTCCTCTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCA 3690
              ||||:|||||||||||||||| ||||||||| |||||| ||||||||||||||||||||
Db        938 GGGCYTAGCTCACCAAGTCCTCTCGCTGTGGAACAGCCAACGGCAGTTAATATTTCCGCA 879

Qy       3691 AATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGG 3750
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        878 AATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGG 819

Qy       3751 TCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTA 3810
              |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db        818 TCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTCGCTGTA 759

Qy       3811 AAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGC 3870
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        758 AAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTGTGAACAAGCTTTTCGGGC 699

Qy       3871 CCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCT 3930
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        698 CCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCT 639

Qy       3931 TCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTC 3990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 TCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTC 579

Qy       3991 AACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGA 4050
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db        578 AACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATNCCTGCTGTGAGA 519

Qy       4051 AACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCA 4110
              ||||||||||:||||||||||||||| |||||||||||||||||||||| ||||||||||
Db        518 AACTTTAAAGTYTCAAATACCCAAGAGGCTTCAGTGTCCATAGTGGATAACTATGAGCCA 459

Qy       4111 AGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGC 4170
              |||||||||||||||||||| |||||||||||||||||: ||| || ||||||||||| ||
Db        458 AGGAGACAGGCGGTGAGAAGTAACAACTCTGAAGTGAMACTGCCCCCCTGTGACCTTCGC 399

Qy       4171 AGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCT 4230
              |||||| ||||:|| :||| ||||||||||||||||||||: |||:|| ||||||||||||:
Db        398 AGTGATGMCCAGVGCMSCCGMCCTTGTGAGGATGGAGCVTCAVGCMCCCATCATCACTCR 339
```

```
Qy       4231 TCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||||| ||||||||||| |||||||||||||||||||||||
Db        338 TCAGTCATTTTTATTTTCTGTCTCAAGCTTCTGAACTTTATGGAACTTTGGCTGTGA 282


RESULT 9
BX386456/c
LOCUS       BX386456                 988 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  BX386456 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 3-PRIME, mRNA sequence.
ACCESSION   BX386456
VERSION     BX386456.1  GI:30453437
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 988)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS1AI021ZB10NP1&cluster=3578.f. Contact :
            Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS1AI021ZB10NP1.
FEATURES             Location/Qualifiers
     source          1. .988
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      384 a    133 c    164 g    249 t     58 others
ORIGIN

  Query Match             14.2%;  Score 608.6;  DB 13;  Length 988;
  Best Local Similarity   90.6%;  Pred. No. 3.6e-124;
  Matches  642;  Conservative   12;  Mismatches   53;  Indels    2;  Gaps    2;

Qy       3580 TTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGC 3639
              |||||||||||||||||||||||||||||||||||||||||||||||:||||| ||||||
Db        988 TTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTRACGGGCCCTAGC 929

Qy       3640 TCACCAAGTCCTCTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTT 3699
              ||||||||||||||||||||||||||||||| |||||||||||||:||||||||||||||
Db        928 TCACCAAGTCCTCTTGCTGTGGTACAGCCAACGGCAGTTAATATTKCCGCAAATGGTTTT 869

Qy       3700 GGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGA 3759
              ||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||||
Db        868 GGATTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTSTGGGTCTTCTAGA 809

Qy       3760 AGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAAT 3819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        808 AGACGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAAT 749

Qy       3820 AAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGG 3879
              |||||||||||||||||||||||||||||||||||||||| ||||||   ||| | ||  
Db        748 AAAGATGATCTCAATCATGTGGATTTGAATGTGTGTACAATCTTTTCNNNCCC-GNTANT 690

Qy       3880 AGTGGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCA 3939
              | |  |||| |||||||  || |||||||||||| | |||||||  |||||||| || ||
Db        689 ATTTNCATGNCTCTTATNNAATTTAACCTATTAANTNGCTTTATNNTGCCTTCANAANCA 630

Qy       3940 ATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTAT 3999
              |||||||||| |||||||||||| |||||||||||||:||||||||||||||||||||||
Db        629 ATTTCTCTGANCGAGACAGTGAATAAAGTGGAATATGWTCATGGAAAACTCAACCTCTAT 570

Qy       4000 TTAGATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAA 4059
              ||||||||| ||||||||||||: |||| ||||||||||||| |||||||||||||||||
Db        569 TTAGATTCTTTAAATGAAACCCWTTTTTTTGTTAATATTCCTNCTGTGAGAAACTTTAAA 510

Qy       4060 GTTTCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAG 4119
              ||||||||||||||| || |||||      ||| ||||||||||||||||||||| ||||
Db        509 GTTTCAAATACCCAANATNCTTCAATNTAAATANTGGATTACTATGAGCCAAGGAMACAG 450

Qy       4120 GCGGTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTC 4179
               || |||||||||||||||||||| | ||||||:|||||| | |||||| || | || :|
Db        449 NCGNTGAGAAGTTACAACTCTGAAATAAAGCTGYCCTCCTTTCACCTTTNCATTCATAHC 390

Qy       4180 CAGGGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTC-CCATCATCACTCTTCAGTCAT 4238
              ||| ||: |||||||| |||| |||||||:||| ||:| |||||||||||||||||||||
Db        389 CAGCGCHCCCGTCCTTTTGAGTATGGAGCHTCAMGCYCACCATCATCACTCTTCAGTCAT 330

Qy       4239 TTTTATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||||:|||||||||||||||||||| || ||:|
Db        329 TTTTATTTTCTGTTTCAAGCKTCTGTACTTTATGGAACTTTGACTTTKA 281


RESULT 10
BU624356/c
LOCUS       BU624356                 638 bp    mRNA    linear   EST 23-SEP-2002
DEFINITION  UI-H-FG1-bgj-l-22-0-UI.s1 NCI_CGAP_FG1 Homo sapiens cDNA clone
            UI-H-FG1-bgj-l-22-0-UI 3', mRNA sequence.
ACCESSION   BU624356
VERSION     BU624356.1  GI:23290571
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 638)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Martin
             cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
             cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
             DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
             Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            Seq primer: M13 FORWARD
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FG1-bgj-l-22-0-UI"
                     /tissue_type="Cell lines"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI_CGAP_FG1"
                     /note="Organ: Enchondroma; Vector: pT7T3-Pac (Pharmacia)
                     with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
                     NCI_CGAP_FG1 is a normalized cDNA library obtained from a
                     pool of mRNA from 2 cell lines from Enchondroma tissues.
                     The library was constructed according to Bonaldo, Lennon
                     and Soares, Genome Research, 6:791-806, 1996. First strand
                     cDNA synthesis was primed with an oligo-dT primer
                     containing a Not I site.  Double stranded cDNA was ligated
```

```
            to an EcoR I adaptor, digested with Not I, and cloned
            directionally into pT7T3-Pac vector. The oligonucleotide
            used to prime the synthesis of first-strand cDNA contains
            a library tag sequence that is located between the Not I
            site and the (dT)18 tail. The sequence tag for this
            library is CGGTCACTC. The cell lines were provided by Dr.
            James Martin from the University of Iowa.
            TAG_LIB=UI-H-FG1
            TAG_TISSUE=Enchondroma cell line (Mix of EN1 and EN2)
            TAG_SEQ=CGGTCACTC"
BASE COUNT      203 a    139 c    109 g    187 t
ORIGIN

  Query Match             13.2%;  Score 565.4;  DB 13;  Length 638;
  Best Local Similarity   99.5%;  Pred. No. 1.3e-114;
  Matches  578;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy       3707 CTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAA 3766
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        638 CTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAA 579

Qy       3767 GATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATG 3826
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        578 GATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATG 519

Qy       3827 ATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCA 3886
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        518 ATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCA 459

Qy       3887 TGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTC 3946
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        458 TGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTC 399

Qy       3947 TGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATT 4006
              ||||||||  | ||||||||||||||||||||||||||||||||||||||||||||||||
Db        398 TGAGCGAG--ATTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATT 341

Qy       4007 CTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAA 4066
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 CTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAA 281

Qy       4067 ATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGA 4126
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 ATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGA 221

Qy       4127 GAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCT 4186
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 GAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCT 161

Qy       4187 GCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTT 4246
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 GCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTT 101

Qy       4247 TCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 4287
              |||||||||||||||||||||||||||||||||||||||||
Db        100 TCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 60


RESULT 11
AA418644
LOCUS       AA418644                 634 bp    mRNA    linear   EST 16-OCT-1997
DEFINITION  zv93f07.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:767365 5'
            similar to TR:G534873 G534873 ALPHA-2-MACROGLOBULIN. ;, mRNA
            sequence.
ACCESSION   AA418644
VERSION     AA418644.1  GI:2080463
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 634)
  AUTHORS   Hillier,L.; Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 494.
FEATURES             Location/Qualifiers
     source          1..634
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:767365"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhHMPu_S1"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      174 a    109 c    139 g    212 t
ORIGIN

  Query Match             12.2%;  Score 525;  DB 9;  Length 634;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-105;
  Matches  525;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3763 CGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAA 3822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAA 60

Qy       3823 GATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGT 3882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGT 120

Qy       3883 GGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATT 3942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATT 180

Qy       3943 TCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTA 4002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTA 240

Qy       4003 GATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTT 4062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTT 300

Qy       4063 TCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCG 4122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCG 360

Qy       4123 GTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAG 4182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAG 420

Qy       4183 GGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTT 4242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        421 GGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTT 480

Qy       4243 ATTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 4287
              |||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 525


RESULT 12
BB613975
LOCUS       BB613975                 648 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB613975 RIKEN full-length enriched, 0 day neonate head Mus
            musculus cDNA clone 4832419I01 5', mRNA sequence.
ACCESSION   BB613975
VERSION     BB613975.1  GI:16454440
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 648)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1..648
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4832419I01"
                     /sex="mixed"
                     /tissue_type="head"
                     /dev_stage="0 day neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 0 day neonate
                     head"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 100.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC
                     3']. cDNA was cloned into the XhoI and BamHI sites.
                     Vector: a modified pBluescript KS(+) after bulk excision
                     from Lambda FLC I."
BASE COUNT      195 a    118 c    140 g    194 t      1 others
ORIGIN

  Query Match            11.7%;  Score 501.6;  DB 10;  Length 648;
  Best Local Similarity   86.0%;  Pred. No. 2e-100;
  Matches  555;  Conservative    0;  Mismatches   90;  Indels    0;  Gaps    0;

Qy       2585 AATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGTT 2644
              ||||  |||||| ||  |||| || |  |    || | ||||  | | || || |||| |
Db          3 AATCTGTCTTATTGGATCTGACAGATAGCAACGTAGAAAGTAAGCAGCAATCTATGAGAT 62

Qy       2645 TCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGAG 2704
              |||| |||||||| |||| || |  |||||||||||||||||||||| |||||||||
Db         63 TCTCCTTTCCTCCTGATACGGTCATTGGCAGTGAAAGAGTTCAGATCACAGCAATTGGAG 122

Qy       2705 ATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGTG 2764
              || | |||||| | ||||||||||||||  | ||| |||| ||||||||||||| ||||
Db        123 ATATCCTTGGTTCCTCCATCAATGGCTTATCTTCACTGATCCGGATGCCTTATGGATGTG 182

Qy       2765 GTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAAA 2824
              ||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db        183 GTGAACAGAACATGATATATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAAC 242

Qy       2825 AGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTACC 2884
              |||||| ||||||| |||||| |||||||| || ||||||| ||||||||||||||||
Db        243 AGAAACAGCTGACAGTTAATTTAAAAGAAAAGGCCCTTTCATATATGAGGCAAGGTTACC 302

Qy       2885 AGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGACC 2944
              | || || |||||||||||||||||||||||| ||||||||||||||| |    ||||
Db        303 AAAGGGAGCTTCTCTATCAGAGGGAAGATGGCTCCTTCAGTGCTTTTGGGGACATTGACT 362

Qy       2945 CTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCTT 3004
              |||||||||||||||| ||||||| ||||||||||||| || || ||||| |||    |
Db        363 CTTCTGGGAGCACTTGGCTGTCAGCATTTGTTTTAAGATGCTTTCTGGAAGCTGATTACT 422

Qy       3005 ACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAGA 3064
              | ||||||||||||||| ||||||||||||||||||| ||||||||||| | |||| |||
Db        423 ATATAGATATTGATCAGGATGTGTTACACAGAACATATACTTGGCTTAATGCACATAAGA 482

Qy       3065 AATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGCA 3124
              ||| ||| ||||||||||||| |||||||||||||||| ||||| ||||||||||||
Db        483 AATTCAATGGTGAATTTTGGGAGCCAGGAAGAGTGATTCACAGTGAACTTCAAGGTGGCA 542

Qy       3125 ATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAGT 3184
                ||||| ||||||||||||| |||||||||| |||||| |||||||||    |||||
Db        543 CCAAAAGCCCAGTAACACTTACGGCCTATATTGTGACTTCTGTCCTGGGATANCAAAAGT 602

Qy       3185 ATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGT 3229
              ||||||||| || ||||| ||||| || ||| | |||||||| |
Db        603 ATCAGCCTAATATCGATGTACAAGACTCAATCAAGTTTTTGGAAT 647


RESULT 13
```

```
CB158337
LOCUS       CB158337                613 bp    mRNA    linear   EST 29-JAN-2003
DEFINITION  K-EST0217654 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-10-H06
            5', mRNA sequence.
ACCESSION   CB158337
VERSION     CB158337.1  GI:28143473
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 613)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 10  row: H  column: 06
            High quality sequence stop: 613.
FEATURES             Location/Qualifiers
     source          1..613
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="L18POOL1n1-10-H06"
                     /cell_line="SNU-354+Cho-CK+Choi-CK+HLK-3"
                     /lab_host="Top10F'"
                     /clone_lib="L18POOL1n1"
                     /note="Organ: Liver; Vector: pT7T3-Pac; Site_1: EcoRI;
                     Site_2: NotI; The library was contributed by the Soares
                     laboratory and it was constructed as described by Bonaldo,
                     M.F., Lennon, G. and Soares, M.B. (1996), Genome Research
                     6(9): 791-806. RNA was prepared from harvested cell
                     culture."
BASE COUNT      183 a    111 c    142 g    177 t
ORIGIN

  Query Match             11.6%;  Score 498;  DB 14;  Length 613;
  Best Local Similarity   99.8%;  Pred. No. 1.2e-99;
  Matches  509;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       3652 CTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATT 3711
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        104 CTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATT 163

Qy       3712 TGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCT 3771
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        164 TGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCT 223

Qy       3772 ATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTC 3831
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        224 ATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTC 283

Qy       3832 AATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCT 3891
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        284 AATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCT 343

Qy       3892 CTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGC 3951
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        344 CTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGC 403

Qy       3952 GAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTA 4011
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        404 GAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTA 463

Qy       4012 AATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACC 4071
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        464 AATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACC 523

Qy       4072 CAAG-ATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAG 4130
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        524 CAAGAATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAG 583

Qy       4131 TTACAACTCTGAAGTGAAGCTGTCCTCCTG 4160
              ||||||||||||||||||||||||||||||
Db        584 TTACAACTCTGAAGTGAAGCTGTCCTCCTG 613


RESULT 14
BX472602
LOCUS       BX472602                515 bp    mRNA    linear   EST 12-JUN-2003
DEFINITION  DKFZp686C02145_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686C02145 5', mRNA sequence.
ACCESSION   BX472602
VERSION     BX472602.1  GI:31666864
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 515)
  AUTHORS   Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
            Fobo,G., Han,M. and Wiemann,S.
  TITLE     EST (Bloecker,H., Boecher,M., Mewes,H.W., Weil,B., Amid,C., et al.)
  JOURNAL   Unpublished
COMMENT     Contact: Bloecker H
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            s1 sequence also available.
            This clone (DKFZp686C02145) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1..515
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686C02145"
                     /tissue_type="human skeletal muscle"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="686 (synonym: hlcc3)"
                     /note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
                     cDNA-collection"
BASE COUNT      130 a    142 c    120 g    123 t
ORIGIN

  Query Match             11.6%;  Score 496;  DB 13;  Length 515;
  Best Local Similarity   100.0%;  Pred. No. 3.3e-99;
  Matches  496;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         20 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 79

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 139

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 199
```

```
Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 259

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        260 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 319

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        320 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 379

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        380 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 439

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        440 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 499

Qy        481 AAAACCTCTTTAAACA 496
              ||||||||||||||||
Db        500 AAAACCTCTTTAAACA 515


RESULT 15
BX370169
LOCUS       BX370169                 919 bp    mRNA    linear   EST 08-MAY-2003
DEFINITION  BX370169 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CS0DI081YB18 5-PRIME, mRNA sequence.
ACCESSION   BX370169
VERSION     BX370169.1  GI:30455906
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 919)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            Library was constructed by Life Technologies, a division of
            Invitrogen. This sequence belongs to sequence cluster 3578.f For
            more information about this cluster, see
            http://www.genoscope.cns.fr/
            cgi-bin/cluster.cgi?seq=CS0BAG038ZH04_CS03584_1&cluster=3578.f.
            Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
            Faraday Avenue Genoscope sequence ID : CS0BAG038ZH04_CS03584_1.
FEATURES             Location/Qualifiers
     source          1..919
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="CS0DI081YB18"
                     /tissue_type="PLACENTA COT 25-NORMALIZED"
                     /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                     /note="1st strand cDNA was primed with a NotI-oligo(dT)
                     primer. Five prime end enriched, double-strand cDNA was
                     digested with Not I and cloned into the Not I and EcoR V
                     sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      285 a    213 c    177 g    242 t      2 others
ORIGIN

  Query Match             11.5%;  Score 493.8;  DB 13;  Length 919;
  Best Local Similarity   94.4%;  Pred. No. 1.1e-98;
  Matches  577;  Conservative    0;  Mismatches   27;  Indels    7;  Gaps    6;

Qy        722 TGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTGA 781
              |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db          4 TGAATTCTAAGCCTTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTGA 63

Qy        782 AAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATTA 841
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         64 AAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATTA 123

Qy        842 CAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAA 901
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        124 CCAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAA 183

Qy        902 ATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA- 960
              |||||||||||||||||||||||||||||||||||| ||||||||||| |||||||||| 
Db        184 ATGTAATGGATTCTTCAAATGGACTTTCTGAATACCCGGATCTATCTTTCCCTGGACCAC 243

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||| ||||| ||||||||||| ||||||||||||||||||||||||
Db        244 GTAGAAATTTTAACCACCGTGACCGAATCAGTTACCGGTATTTCAAGAAATGTAAGCACT 303

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTA-CTGTCTT 1079
              |||||||||||||||| ||||||||||||||||||| |||||||||||||| ||| |||
Db        304 AATGTGTTCTTCAAGCCACATGATTACATCATTGAGGTTTTGATTATACTACCTGGCTT 363

Qy       1080 GAAGCCATCTCTCAACTTCACAGCCA-CTGTGAAGGTAACTCGTGCTGATGGCAACCAAC 1138
              |||||||||||||||||| ||||||| |||||||||||||||||||||||||||||| ||
Db        364 GAAGCCATCTCTCAACTTTACAGCCACCTGTGAAGGTAACTCGTGCTGATGGCAACCCAC 423

Qy       1139 TGACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTG 1198
              |||||||||||||||||||||||||||||||| ||||||||||||| |||||||||||||
Db        424 TGACTCTTGAAGAAAGAAGAAATAATGTAGTCCTAACAGTGACACACAGAAACTATACTG 483

Qy       1199 AGTACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATT 1258
              ||||||||||||||||||||||||| |||||||||| |||||||||||| |||||||| |
Db        484 AGTACTGGAGCGGATCTAACAGTGGGAATCAGAAAAAGGAAGCTGTTCACAAAATAAAAT 543

Qy       1259 A-TACTGTCCCCCAAAGTGGAACTT-TTAAGATTGAATT--CCCAATCCTGGAGGATTCC 1314
              | |||||||||||||||||| |||| |||||||||||||  ||||  |   |||||||||
Db        544 ATTACTGTCCCCCAAAGTGGGACTTCTTAAGATTGAATTTCCCCACCCTGCGAGGATTCC 603

Qy       1315 AGTGAGCTACA 1325
              || ||||  ||
Db        604 AGCGAGCCTCA 614

Search completed: January 16, 2004, 20:57:23
Job time : 7735 secs
```

```
OM nucleic - nucleic search, using sw model

Run on:          January 16, 2004, 12:38:20 ; Search time 14986 Seconds
                                              (without alignments)
                                              11702.894 Million cell updates/sec

Title:           US-10-020-095-3
Perfect score:   4287
Sequence:        1 atgcagggcccaccgctcct..........ttatggaactttggctgtga 4287

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters:      5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*
                 15:  em_ba:*
                 16:  em_fun:*
                 17:  em_hum:*
                 18:  em_in:*
                 19:  em_mu:*
                 20:  em_om:*
                 21:  em_or:*
                 22:  em_ov:*
                 23:  em_pat:*
                 24:  em_ph:*
                 25:  em_pl:*
                 26:  em_ro:*
                 27:  em_sts:*
                 28:  em_un:*
                 29:  em_vi:*
                 30:  em_htg_hum:*
                 31:  em_htg_inv:*
                 32:  em_htg_other:*
                 33:  em_htg_mus:*
                 34:  em_htg_pln:*
                 35:  em_htg_rod:*
                 36:  em_htg_mam:*
                 37:  em_htg_vrt:*
                 38:  em_sy:*
                 39:  em_htgo_hum:*
                 40:  em_htgo_mus:*
                 41:  em_htgo_other:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	4224.4	98.5	4761	6	AX534934	AX534934 Sequence
	2	4224.4	98.5	5883	9	AF410459	AF410459 Homo sapi
	3	4224.4	98.5	5895	6	AX534938	AX534938 Sequence
	4	4222.8	98.5	4761	6	AX534936	AX534936 Sequence
	5	4222.8	98.5	5895	6	AX534940	AX534940 Sequence
	6	4221.4	98.5	4335	6	AX537610	AX537610 Sequence
	7	4219.8	98.4	4335	6	AX537612	AX537612 Sequence
	8	3187.4	74.4	3535	6	AX534942	AX534942 Sequence
	9	3185.8	74.3	3535	6	AX534944	AX534944 Sequence
	10	2709.2	63.2	4735	10	AY083458	AY083458 Mus muscu
	11	2709.2	63.2	5644	10	BC052443	BC052443 Mus muscu
	12	2448.2	57.1	2938	9	HSM805581	AL834478 Homo sapi
	13	1969.4	45.9	2273	9	AK095888	AK095888 Homo sapi
	14	730.8	17.0	1300	6	AX083146	AX083146 Sequence
	15	232	5.4	163577	9	AL590428	AL590428 Human DNA
	16	230	5.4	176144	2	AC012408	AC012408 Homo sapi
	17	219.6	5.1	5603	3	CIN431688	AJ431688 Ciona int
	18	203.8	4.8	4678	3	D83196	D83196 Limulus sp.
c	19	182	4.2	251250	2	AC097023	AC097023 Rattus no
	20	177.4	4.1	91419	9	AL591480	AL591480 Human DNA
	21	175	4.1	149597	2	AC034271	AC034271 Homo sapi
	22	174	4.1	4527	6	AX054953	AX054953 Sequence
	23	174	4.1	4560	6	AX054954	AX054954 Sequence
	24	171.6	4.0	2629	3	AY118302	AY118302 Drosophil
	25	167.6	3.9	4715	5	GGOVOM	X78801 G.gallus mR
	26	167	3.9	4578	3	AY122084	AY122084 Drosophil
	27	167	3.9	4579	3	DME269539	AJ269539 Drosophil
c	28	160	3.7	144589	2	AC112668	AC112668 Mus muscu
	29	160	3.7	191258	2	AC118254	AC118254 Mus muscu
	30	155.4	3.6	4620	10	RNA1I3	X52984 Rat mRNA fo
c	31	154.8	3.6	176144	2	AC012408	AC012408 Homo sapi
	32	154.2	3.6	4636	10	RATA1I3A	J03552 Rat plasma
	33	152.2	3.6	4220	3	DME269538	AJ269538 Drosophil
	34	151.2	3.5	4575	10	GPIMSPB	D84339 Cavia porce
	35	149.8	3.5	2501	10	RATA1I3	J03524 Rat alpha-1
	36	146.4	3.4	4950	5	XELENDO	L63543 Xenopus lae
c	37	143.8	3.4	122013	3	AC092399	AC092399 Drosophil
c	38	143.8	3.4	164414	3	AC092395	AC092395 Drosophil
c	39	143.8	3.4	168503	3	AC115484	AC115484 Drosophil
c	40	143.8	3.4	275699	3	AE003649	AE003649 Drosophil
c	41	143.8	3.4	337178	3	DROSADH08	AE003414 Drosophil
c	42	140.8	3.3	191258	2	AC118254	AC118254 Mus muscu
	43	136	3.2	4623	10	BC036299	BC036299 Mus muscu
	44	136	3.2	4672	10	BC051037	BC051037 Mus muscu
	45	134.8	3.1	2689	5	AB026130	AB026130 Cyprinus

ALIGNMENTS

```
RESULT 1
AX534934
LOCUS       AX534934                4761 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 1 from Patent WO02070696.
ACCESSION   AX534934
VERSION     AX534934.1  GI:25261474
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 1 12-SEP-2002;
```

```
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1..4761
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113..4450
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57231.1"
                     /db_xref="GI:25261475"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1420 a    912 c   1016 g   1413 t
ORIGIN

  Query Match             98.5%;  Score 4224.4;  DB 6;  Length 4761;
  Best Local Similarity   98.8%;  Pred. No. 0;              4450
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy   1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy   1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy   1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy   1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy   1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy   1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy   1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy   1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy   2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy   2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy   2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy   2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy   2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy   2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy   2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy   2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy   3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                    ||||||||||||||||||
Db   3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832
```

```
Qy        3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy        3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy        3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy        3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy        3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy        3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy        4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy        4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy        4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy        4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy        4270 ATGGAACTTTGGCTGTGA 4287
               ||||||||||||||||||
Db        4433 ATGGAACTTTGGCTGTGA 4450

RESULT 2
AF410459
LOCUS       AF410459                5883 bp    mRNA    linear   PRI 02-MAR-2002
DEFINITION  Homo sapiens CD109 (CD109) mRNA, complete cds.
ACCESSION   AF410459
VERSION     AF410459.1  GI:19071208
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5883)
  AUTHORS   Lin,M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
            Wu,X.F. and Schuh,A.C.
  TITLE     Cell surface antigen CD109 is a novel member of the alpha(2)
            macroglobulin/C3, C4, C5 family of thioester-containing proteins
  JOURNAL   Blood 99 (5), 1683-1691 (2002)
  MEDLINE   21849742
   PUBMED   11861284
REFERENCE   2  (bases 1 to 5883)
  AUTHORS   Lin,M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
            Wu,X.-F. and Schuh,A.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's
            College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada
FEATURES             Location/Qualifiers
     source          1..5883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="K1"
     gene            1..5883
                     /gene="CD109"
     5'UTR           1..112
                     /gene="CD109"
     CDS             113..4450
                     /gene="CD109"
                     /note="associated with the Gov alloantigen system"
                     /codon_start=1
                     /product="CD109"
                     /protein_id="AAL84159.1"
                     /db_xref="GI:19071209"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
     3'UTR           4451..5883
                     /gene="CD109"
     misc_feature    4748
                     /gene="CD109"
                     /note="alternative polyA site found on clone K1"
BASE COUNT     1697 a   1173 c   1241 g   1772 t
ORIGIN

  Query Match             98.5%;  Score 4224.4;  DB 9;  Length 5883;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy           1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy          61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy         121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy         181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy         241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy         301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632
```

```
Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2633 GCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy   3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                    ||||||||||||||||||
Db   3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy   3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy   3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy   3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy   3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy   3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy   3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy   4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy   4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy   4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy   4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy   4270 ATGGAACTTTGGCTGTGA 4287
          ||||||||||||||||||
Db   4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 3
AX534938
LOCUS       AX534938                5895 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 5 from Patent WO02070696.
ACCESSION   AX534938
VERSION     AX534938.1  GI:25261482
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 5 12-SEP-2002;
            Schuh, Andre (CA) , Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1. .5895
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
```

Same as Result 1

```
                     /db_xref="taxon:9606"
     CDS             113..4450
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57233.1"
                     /db_xref="GI:25261483"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1709 a   1173 c   1241 g   1772 t
ORIGIN

  Query Match            98.5%;  Score 4224.4;  DB 6;  Length 5895;
  Best Local Similarity  98.8%;  Pred. No. 0;
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732
```

```
Qy   1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy   1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy   1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy   1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy   1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy   1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy   1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy   2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy   2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy   2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy   2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy   2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy   2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy   2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy   2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy   2761 TGTGGTGAACAGAACATGATAAATTTGCTCCAAATATTTACATTTGGATTATCTGACT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2873 TGTGGTGAACAGAACATGATAAATTTGCTCCAAATATTTACATTTGGATTATCTGACT 2932

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
          |||||||||||||||||||||||||||||||||||||||||||||||||||         
Db   3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy   3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                    ||||||||||||||||||
Db   3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy   3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy   3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4433 ATGGAACTTTGGCTGTGA 4450


RESULT 4
AX534936
LOCUS       AX534936                4761 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 3 from Patent WO02070696.
ACCESSION   AX534936
VERSION     AX534936.1  GI:25261478
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 3 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1..4761
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113..4450
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57232.1"
                     /db_xref="GI:25261479"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1419 a    913 c   1016 g   1413 t
ORIGIN

  Query Match             98.5%;  Score 4222.8;  DB 6;  Length 4761;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4285;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832
```

Same as Result 1

```
Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 5
AX534940
LOCUS       AX534940                5895 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 7 from Patent WO02070696.
ACCESSION   AX534940
VERSION     AX534940.1  GI:25261486
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 7 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1..5895
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113..4450
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57234.1"
                     /db_xref="GI:25261487"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
```

Same as Result 1

```
                    YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                    IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                    VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                    LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                    VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                    LLYFMELWL"
BASE COUNT     1708 a   1174 c   1241 g   1772 t
ORIGIN

  Query Match             98.5%;  Score 4222.8;  DB 6;  Length 5895;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4285;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
```

```
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTGCTATTTGTCAGCTCAATGTTGTA 3729
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312
```

```
Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4433 ATGGAACTTTGGCTGTGA 4450


RESULT 6
AX537610
LOCUS       AX537610                4335 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO02070738.
ACCESSION   AX537610
VERSION     AX537610.1  GI:25269531
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Ouwehand,W.
  TITLE     Diagnosis and treatment of blood disorders
  JOURNAL   Patent: WO 02070738-A 1 12-SEP-2002;
            Schuh, Andre (CA)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..>4335
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57496.1"
                     /db_xref="GI:25269532"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1296 a    837 c    922 g   1280 t
ORIGIN

  Query Match             98.5%;  Score 4221.4;  DB 6;  Length 4335;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4283;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy         1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy         1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy         1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy         1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy         1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy         1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy         1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy         1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy         1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy         1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy         1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy         1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy         1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy         1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy         1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy         2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy         2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy         2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy         2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy         2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy         2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
                ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db         2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400

Qy         2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy         2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy         2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy         2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy         2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy         2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy         2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy         2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy         2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy         2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy         3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy         3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy         3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy         3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy         3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
```

```
Qy      3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy      3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy      3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy      3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy      3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy      3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy      3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                       ||||||||||||||||||
Db      3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy      3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy      3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy      3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy      3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy      3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy      3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy      4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy      4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy      4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy      4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy      4270 ATGGAACTTTGGCTG 4284
             |||||||||||||||
Db      4321 ATGGAACTTTGGCTG 4335
```

```
RESULT 7
AX537612
LOCUS       AX537612                4335 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 3 from Patent WO02070738.
ACCESSION   AX537612
VERSION     AX537612.1  GI:25269535
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Ouwehand,W.
  TITLE     Diagnosis and treatment of blood disorders
  JOURNAL   Patent: WO 02070738-A 3 12-SEP-2002;
            Schuh, Andre (CA)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..>4335
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57497.1"
                     /db_xref="GI:25269536"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1295 a    838 c    922 g   1280 t
ORIGIN
```

SAME AS Result 6

```
  Query Match             98.4%;  Score 4219.8;  DB 6;  Length 4335;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4282;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400
```

```
Qy        2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy        2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy        2521 GCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2521 GCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy        2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy        2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy        2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy        2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy        2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy        2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy        2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy        3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy        3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy        3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy        3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy        3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300

Qy        3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy        3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy        3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy        3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy        3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy        3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy        3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                         ||||||||||||||||||
Db        3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy        3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy        3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy        3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy        3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy        3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy        3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy        4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy        4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy        4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy        4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4269
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy        4270 ATGGAACTTTGGCTG 4284
               |||||||||||||||
Db        4321 ATGGAACTTTGGCTG 4335
```

```
RESULT 8
AX534942
LOCUS       AX534942                3535 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 9 from Patent WO02070696.
ACCESSION   AX534942
VERSION     AX534942.1  GI:25261490
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
```

```
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 9 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1..3535
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113..3316
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57235.1"
                     /db_xref="GI:25261491"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQVFRI"
BASE COUNT     1108 a    675 c    731 g   1021 t
ORIGIN

  Query Match            74.4%;  Score 3187.4;  DB 6;  Length 3535;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 3188;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       3181 AAGTATCAG 3189
              |||||||||
Db       3293 AAGTATCAG 3301
```

```
RESULT 9
AX534944
LOCUS       AX534944                3535 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 11 from Patent WO02070696.
ACCESSION   AX534944
VERSION     AX534944.1  GI:25261494
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 11 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1..3535
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             113..3316
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57236.1"
                     /db_xref="GI:25261495"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
```

```
                  VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                  KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                  MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                  SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                  ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                  INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                  EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                  IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                  SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                  GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQVFRI"
BASE COUNT     1107 a    676 c    731 g   1021 t
ORIGIN

  Query Match             74.3%;  Score 3185.8;  DB 6;  Length 3535;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3187;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952
```

```
Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       3181 AAGTATCAG 3189
              |||||||||
Db       3293 AAGTATCAG 3301
```

```
RESULT 10
AY083458
LOCUS       AY083458                4735 bp    mRNA    linear   ROD 08-APR-2002
DEFINITION  Mus musculus GPI-anchored alpha-2 macroglobulin-related protein
            mRNA, complete cds.
ACCESSION   AY083458
VERSION     AY083458.1  GI:20070079
KEYWORDS    .
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4735)
  AUTHORS   Hashimoto,M., Ichihara,M. and Takahashi,M.
  TITLE     Cloning and characterization of GPI-anchored alpha-2
            macroglobulin-related protein
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4735)
  AUTHORS   Hashimoto,M., Ichihara,M. and Takahashi,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-2002) Department of Pathology, Nagoya University
            Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya,
            Aichi 466-8550, Japan
FEATURES             Location/Qualifiers
     source          1..4735
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /cell_line="NIH3T3"
     CDS             343..4671
                     /note="GARP"
                     /codon_start=1
                     /product="GPI-anchored alpha-2 macroglobulin-related
                     protein"
                     /protein_id="AAM00021.1"
                     /db_xref="GI:20070080"
                     /translation="MRSRRLLSAAHLLCLCAVALAAPGSRFLVTAPGIIRPGANVTIG
                     VDLLENSPPQVLVKAQVFKIASNKSRSILEAEGVFHRGHFKTLVLPALPLSSADKIYE
                     LHINGQSENEIVFSNRTRLTFESKSISVLIQTDKAFYKPKQEVKFRVLTLCSDLKPYR
                     TSVDIFIKDPKSNVIQQWFSQKGDLGVVSKTFQLSSNPIFGDWSIQVQVNDQQYYQSF
                     QVLEYVLPKFEVTVQTPLYCSLKSKQLNGSVIAKYTYGKPVKGSLSLTFLPLSFWGKK
                     KNITKSFEINGFANFSFDNYEMKKVMNLKPLTDVSEGSYENVDPSFPGPAEIIATVTE
                     SLTGISRMASTNVFFKQHDYIIEIFDYTTVLKPSLNFTATVKVSRSDGNQLTPEEIEN
                     DLVTVVTQRKNNHPESQRDQEMDYIQTVNYTIPQNGIIKIEFPVMSISGELQLKAYFL
                     DGTSSVTVHSMFTSPSKTYIQLKTRDEYIKVGSPFDLMVSGNRQFKDLSYMVISKGQL
                     VAAGKQSSRTFSLTPEASWAPKACIIAYYIAEDGEIINDILKIPVQLVFENKVKLFWS
                     KPTVKPSDKVSLRISATQSDSLVGIVAVDKSVTLMENSNSITMETMVHELELYNTEYY
                     LGMFMNSFAVFQECGLWVLTDATLIRDSIDEVYDTEEYSERFAEENEANLVDFEDASS
                     VNNVHVRKNFPETWIWLDAYMGSKIYEEFEVTVPDSITSWVASAFVISEDLGFGLTTV
                     PAELQAFQPFFLFLNLPYSVIRGEEFALEVSIVNYLKDTIKVVILIEESDSFDILMTS
                     NDTNGTIYRKTVQVPRDNGVTLVFPIKPTHLGEIPITVTAASPTASDAVTQTIVVKPE
                     GIEKSYSKSVLLDLTDSNVESKQQSMRFSFPPDTVIGSERVQITAIGDILGSSINGLS
                     SLIRMPYGCGEQNMIYFAPNIYILDYLTKQKQLTVNLKEKALSYMRQGYQRELLYQRE
                     DGSFSAFGDIDSSGSTWLSAFVLRCFLEADYYIDIDQDVLHRTYTWLNAHKKFNGEFW
                     EPGRVIHSELQGGTKSPVTLTAYIVTSVLGYKKYQPNIDVQDSIKFLEFEFSRGISDN
                     YTLAIISYALSTVGSPKAEEALNLLMQRSEKEGDTQFWLSSGPALSGSWQPRSVDIEI
                     AAYALLAHTLHHVSEGIPVMRWLIQQRNSLGGFVSTQDTVVALKALSEFSALVHKENT
                     DIQLTVTGPGIPRSIHFRIDSQNLFLLHQEELHALDPITVNVSAHGSGFAICQLNVDY
```

```
                    NVKGSGSSKRRRSTENQEVFDLDVIVNNEDDISHLNLNVCTSHLGSERTGMVLMEVNL
                    LSGFSASSDSIPLSETLKKVEYDNGKLNLYLDSVNESQFCVNIPTVRDYKVSNIRDGS
                    VSVMDYYEPRRQAVRSYNTQVKLSSCYLSPDTNCKSHTDGATDSLRRSSSLLVFCSVL
                    LYFVQH"
BASE COUNT     1338 a   1034 c   1087 g   1276 t
ORIGIN

  Query Match            63.2%;  Score 2709.2;  DB 10;  Length 4735;
  Best Local Similarity  77.9%;  Pred. No. 0;
  Matches 3393;  Conservative    0;  Mismatches  873;  Indels   90;  Gaps    7;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
Db        343 ATGCGCAGTCGGAGGCTCCTGAGCGCGGCCCACCTTCTCTGCTTGTGCGCGGTCGCGCTG 402

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
Db        403 GCC---GCCCCAGGGTCCCGCTTTCTGGTGACAGCCCCAGGGATCATCAGACCCGGAGCA 459

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
Db        460 AATGTGACTATTGGGGTGGATCTCCTGGAAAATAGCCCCCCACAGGTCCTTGTAAAGGCT 519

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
Db        520 CAGGTGTTCAAGATAGCTTCCAACAAATCGAGATCCATCCTGGAAGCAGAAGGAGTCTTT 579

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
Db        580 CACAGAGGCCATTTCAAGACTCTCGTTCTTCCGGCACTACCTCTGAGCAGTGCAGATAAG 639

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
Db        640 ATTTATGAGCTACACATAAACGGACAATCAGAGAATGAGATCGTATTCTCCAACAGGACA 699

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
Db        700 CGTTTAACATTTGAGAGCAAGAGTATATCTGTCCTCATTCAGACAGACAAGGCATTCTAC 759

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
Db        760 AAGCCGAAGCAGGAAGTGAAGTTTCGGGTGCTCACACTCTGCTCAGATCTCAAGCCTTAT 819

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
Db        820 AGGACATCCGTGGACATTTTCATTAAGGATCCGAAGTCTAATGTGATCCAACAGTGGTTT 879

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
Db        880 TCTCAGAAAGGTGATCTTGGAGTCGTTTCCAAAACTTTTCAACTATCCTCTAATCCAATA 939

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
Db        940 TTTGGTGACTGGTCCATTCAGGTTCAAGTGAATGATCAGCAATATTACCAGTCATTTCAG 999

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
Db       1000 GTCCTAGAATATGTGTTACCAAAGTTCGAAGTCACCGTGCAGACCCCGCTGTATTGTTCC 1059

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
Db       1060 TTGAAGTCTAAACAGTTAAATGGCTCCGTCATAGCAAAGTATACATACGGGAAGCCAGTG 1119

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
Db       1120 AAGGGAAGTCTGTCACTTACATTTTTACCTTTGTCTTTTTGGGGAAAGAAGAAGAATATC 1179

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
Db       1180 ACAAAAAGTTTTGAGATAAATGGATTTGCAAACTTCTCTTTTGATAATTATGAGATGAAA 1239

Qy        901 AATGTAAT------------------GGATTCTTCAAATGGACTTTCTGAATACCTGGAT 942
Db       1240 AAGGTAATGAATTTGAAACCGCTAACGGATGTCTCGGAAGGGAGTTATGAGAATGTGGAC 1299

Qy        943 CTATCTTCCCCTGGACCAGTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATT 1002
Db       1300 CCGTCCTTTCCTGGCCGGCAGAGATTATAGCTACTGTGACGGAATCACTCACAGGTATC 1359

Qy       1003 TCAAGAAATGTAAGCACTAATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTT 1062
Db       1360 TCAAGAATGGCAAGCACCAATGTGTTTTTCAAGCAACATGATTACATCATTGAGATTTTT 1419

Qy       1063 GATTATACTACTGTCTTGAAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGT 1122
Db       1420 GATTATACGACTGTCTTGAAACCATCTCTCAACTTCACAGCCACTGTGAAGGTCAGCCGC 1479

Qy       1123 GCTGATGGCAACCAACTGACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACA 1182
Db       1480 TCTGATGGCAATCAACTAACTCCTGAAGAAATAGAAAATGATTTGGTCACAGTAGTGACC 1539

Qy       1183 CAGAGAAACTATACTGAGTACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCT 1242
Db       1540 CAGCGGAAAAATAACCACCCTGAGAG---------TCAGAGAGATCAGGAAATGGATTAT 1590

Qy       1243 GTTCAGAAAATAAATTATACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATC 1302
Db       1591 ATCCAGACCGTAAATTATACTATCCCCCAAAATGGAATCATTAAGATTGAATTCCCAGTC 1650

Qy       1303 CTGGAGGATTCCAGTGAGCTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATG 1362
Db       1651 ATGTCCATTTCCGGCGAGCTACAACTGAAGGCTTATTTTCTTGATGGTACAAGTTCTGTG 1710

Qy       1363 GCAGTTCATAGTCTGTTTAAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGAT 1422
Db       1711 ACAGTCCACAGTATGTTCACCTCTCCTAGTAAGACGTACATCCAGCTTAAAACCAGAGAT 1770

Qy       1423 GAAAATATAAAGGTGGGATCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAG 1482
Db       1771 GAATATATAAAGGTGGGGTCACCTTTTGATTTGATGGTTAGTGGCAACAGACAATTCAAG 1830

Qy       1483 GAGTTAAGCTATATGGTAGTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCA 1542
Db       1831 GACTTAAGCTATATGGTAATATCAAAGGGACAGTTGGTGGCTGCAGGCAAACAAAGTTCA 1890

Qy       1543 ACAATGTTCTCTTTAACACCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTAT 1602
Db       1891 AGAACTTTCTCTTTAACACCAGAGGCTTCCTGGGCTCCAAAGGCCTGTATAATCGCATAT 1950

Qy       1603 TATATTGAAGATGATGGGGAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTT 1662
Db       1951 TATATTGCAGAGGATGGGGAAATTATAAATGATATTCTAAAAATCCCCGTTCAGCTTGTT 2010

Qy       1663 TTTAAAAATAAGATAAAGCTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTC 1722
Db       2011 TTTGAAAATAAGGTAAAGCTGTTTTGGAGTAAACCTACAGTCAAGCCATCTGATAAGGTC 2070

Qy       1723 TCTCTTAGGATCTCTGTGACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAA 1782
Db       2071 TCTCTCAGGATCTCTGCGACACAGTCTGACTCCCTGGTGGGGATTGTAGCTGTTGACAAA 2130

Qy       1783 AGTGTGAATCTGATGAATGCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTG 1842
Db       2131 AGTGTAACATTGATGGAAAACTCCAACAGTATTACAATGGAAACCATGGTCCATGAACTG 2190

Qy       1843 GAACTTTATAACACAGGATATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAG 1902
Db       2191 GAACTTTATAACACAGAATATTATCTAGGCATGTTCATGAACTCCTTCGCTGTATTTCAG 2250

Qy       1903 GAATGTGGACTCTGGGTATTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTT 1962
Db       2251 GAATGTGGTCTCTGGGTGTTGACAGATGCAACCCTTATAAGAGATAGCATTGATGAAGTC 2310

Qy       1963 TATGACAATGCAGAATATGCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGAT 2022
Db       2311 TACGATACTGAAGAGTATTCTGAACGCTTTGCAGAGGAAAATGAGGCAAACCTGGTTGAT 2370
```

```
Qy   2023 ATTCATGACTTTTCTTTGGGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGG 2082
Db   2371 TTTGAAGATGCTTCTTCAGTTAACAATGTACATGTCAGGAAGAATTTTCCAGAAACATGG 2430

Qy   2083 ATTTGGCTAGACACCAACATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCT 2142
Db   2431 ATTTGGCTAGACGCCTACATGGGATCTAAGATTTATGAAGAATTTGAAGTTACTGTACCT 2490

Qy   2143 GATTCTATCACTTCTTGGGTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGA 2202
Db   2491 GACTCCATCACTTCTTGGGTGGCTTCGGCTTTTGTCATCTCAGAAGACCTGGGTTTTGGG 2550

Qy   2203 CTAACAACTACTCCAGTGGAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTT 2262
Db   2551 TTAACAACCGTTCCAGCAGAGCTGCAAGCCTTCCAACCGTTTTTCCTTTTCCTGAATCTT 2610

Qy   2263 CCCTACTCTGTTATCAGAGGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTG 2322
Db   2611 CCGTACTCTGTTATCAGAGGTGAAGAGTTTGCTTTGGAAGTATCCATCGTCAATTATTTG 2670

Qy   2323 AAAGATGCCACTGAGGTTAAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATG 2382
Db   2671 AAAGATACAATTAAGGTTGTGATACTCATTGAGGAAAGTGACAGCTTTGATATTTTAATG 2730

Qy   2383 ACTTCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGAT 2442
Db   2731 ACTTCAAATGACACCAATGGCACCATATACCGGAAGACTGTTCAAGTTCCCAGGGATAAT 2790

Qy   2443 GGGGCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTC 2502
Db   2791 GGGGTCACTCTTGTTTTTCCCATCAAGCCCACGCACTTGGGAGAGATTCCCATCACAGTG 2850

Qy   2503 ACAGCTCTTTCACCCACTGCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAA 2562
Db   2851 ACGGCTGCCTCACCCACTGCCTCTGATGCTGTCACCCAGACAATTGTAGTAAAGCCTGAA 2910

Qy   2563 GGAATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAG 2622
Db   2911 GGGATAGAAAAGTCGTATTCGAAATCTGTCTTATTGGATCTGACAGATAGCAACGTAGAA 2970

Qy   2623 AGTACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGA 2682
Db   2971 AGTAAGCAGCAATCTATGAGATTCTCCTTTCCTCCTGATACGGTCATTGGCAGTGAAAGA 3030

Qy   2683 GTTCAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTG 2742
Db   3031 GTTCAGATCACAGCAATTGGAGATATCCTTGGTTCCTCCATCAATGGCTTATCTTCACTG 3090

Qy   2743 ATTCGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTAC 2802
Db   3091 ATCCGGATGCCTTATGGATGTGGTGAACAGAACATGATATATTTTGCTCCAAATATTTAC 3150

Qy   2803 ATTTTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTT 2862
Db   3151 ATTTTGGATTATCTGACTAAACAGAAACAGCTGACAGTTAATTTAAAAGAAAAGGCCCTT 3210

Qy   2863 TCATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTC 2922
Db   3211 TCATATATGAGGCAAGGTTACCAAAGGGAGCTTCTCTATCAGAGGGAAGATGGCTCCTTC 3270

Qy   2923 AGTGCTTTTGGGAATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGA 2982
Db   3271 AGTGCTTTTGGGGACATTGACTCTTCTGGGAGCACTTGGCTGTCAGCATTTGTTTTAAGA 3330

Qy   2983 TGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGAACATAC 3042
Db   3331 TGCTTTCTGGAAGCTGATTACTATATAGATATTGATCAGGATGTGTTACACAGAACATAT 3390

Qy   3043 ACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATT 3102
Db   3391 ACTTGGCTTAATGCACATAAGAAATTCAATGGTGAATTTTGGGAGCCAGGAAGAGTGATT 3450

Qy   3103 CATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACT 3162
Db   3451 CACAGTGAACTTCAAGGTGGCACCAAAAGCCCAGTAACACTTACGGCCTATATTGTGACT 3510

Qy   3163 TCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTT 3222
Db   3511 TCTGTCCTGGGATACAAAAAGTATCAGCCTAATATCGATGTACAAGACTCAATCAAGTTT 3570

Qy   3223 TTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTAT 3282
Db   3571 TTGGAATTTGAATTCAGCAGAGGAATTTCGGACAATTATACCTTAGCAATTATATCCTAT 3630

Qy   3283 GCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGA 3342
Db   3631 GCCCTGTCCACAGTCGGGAGCCCTAAAGCAGAGGAGGCTTTGAACTTGCTGATGCAGCGA 3690

Qy   3343 GCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCC 3402
Db   3691 TCAGAAAAGGAAGGAGACACACAGTTCTGGTTGTCATCAGGACCTGCGCTCTCTGGGTCC 3750

Qy   3403 TGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTA 3462
Db   3751 TGGCAGCCGCGCTCCGTGGACATCGAAATCGCTGCTTATGCACTGCTGGCGCACACTCTG 3810

Qy   3463 CAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGC 3522
Db   3811 CA---CCATGTGTCTGAGGGAATCCCGGTTATGAGGTGGCTCATCCAGCAAAGAAACAGC 3867

Qy   3523 TTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTT 3582
Db   3868 CTGGGAGGTTTTGTATCCACTCAGGATACTGTTGTGGCCCTGAAGGCATTATCTGAATTT 3927

Qy   3583 GCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCA 3642
Db   3928 TCAGCCCTTGTGCACAAGGAAAATACAGATATCCAACTGACCGTCACAGGGCCCGGCATC 3987

Qy   3643 CCAAGTCCT--------------------------------------------------- 3651
Db   3988 CCCAGATCCATACACTTCCGAATTGACTCCCAGAACCTCTTTCTCCTTCACCAGGAAGAG 4047

Qy   3652 CTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATT 3711
Db   4048 CTTCATGCGCTAGATCCCATCACCGTTAATGTTTCTGCGCATGGCTCGGGATTTGCTATT 4107

Qy   3712 TGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCT 3771
Db   4108 TGCCAGCTTAATGTTGACTATAACGTGAAAGGTTCAGGTTCTTCTAAAAGACGGAGATCT 4167

Qy   3772 ATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTC 3831
Db   4168 ACCGAAAATCAAGAAGTTTTTGATTTAGACGTCATTGTGA---ATAATGAGGACGACATT 4224

Qy   3832 AATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCT 3891
Db   4225 AGTCACCTGAATCTGAATGTGTGCACAAGTCACTTGGGTTCAGAGAGGACAGGCATGGTC 4284

Qy   3892 CTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGC 3951
Db   4285 CTTATGGAAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATTCAATTCCTCTGAGT 4344

Qy   3952 GAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTA 4011
Db   4345 GAGACCCTGAAGAAAGTGGAATATGATAACGGGAAACTCAACCTTTATTTAGATTCTGTA 4404

Qy   4012 AATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACC 4071
Db   4405 AATGAATCCCAGTTTTGTGTTAATATTCCCACTGTGAGAGACTACAAAGTTTCAAATATT 4464

Qy   4072 CAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGT 4131
Db   4465 CGAGATGGTTCAGTATCTGTAATGGACTACTATGAACCAAGGAGACAGGCAGTGCGAAGC 4524

Qy   4132 TACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGT 4191
```

```
             |||||| |  | ||||||||||||| ||| ||||   ||||   | |      ||||
Db      4525 TACAACACCCAGGTGAAGCTGTCCTCATGTTACCT---CAGTCCAGACACCAACTGCAAG 4581

Qy      4192 CCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGT 4251
                |     ||| |||||  | | ||||| ||  |  |||||   | || |   | || ||
Db      4582 TCACACACGGACGGAGCCACGGACTCCCTTCGACGTTCTTCGAGCCTTCTCGTCTTTTGC 4641

Qy      4252 TTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 4287
             | |   ||||||||||| || ||| |||  | |||
Db      4642 TCCGTCCTTCTGTACTTTGTGCAACATTGATTATGA 4677


RESULT 11
BC052443
LOCUS       BC052443                5644 bp    mRNA    linear   ROD 19-MAY-2003
DEFINITION  Mus musculus cDNA clone MGC:63463 IMAGE:6406841, complete cds.
ACCESSION   BC052443
VERSION     BC052443.1  GI:30851465
KEYWORDS    MGC.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 5644)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
            Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
  TITLE     Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  MEDLINE   22388257
   PUBMED   12477932
REFERENCE   2  (bases 1 to 5644)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-MAY-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
            Thomas L. Casavant.
            Web site: http://genome.uiowa.edu
            Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
            Bonaldo,M.F., Akabogu,I.,  Bair,T., Bair,J., Crouch,K., Davis,A.,
            Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
            Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,
            Casavant,T., Soares,M.B.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series:  Plate:  Row:  Column: 0.
FEATURES             Location/Qualifiers
     source          1..5644
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="MGC:63463 IMAGE:6406841"
                     /tissue_type="Brain, enriched mouse brain 12.5dp"
                     /clone_lib="NIH_BMAP_FO0"
                     /lab_host="DH10B"
                     /note="Vector: pYX-ASC"
     CDS             105..4433
                     /codon_start=1
                     /product="Unknown (protein for MGC:63463)"
                     /protein_id="AAH52443.1"
                     /db_xref="GI:30851466"
                     /translation="MRSRRLLSAAHLLCLCAVALAAPGSRFLVTAPGIIRPGANVTIG
                     VDLLENSPPQVLVKAQVFKIASNKSRSILEAEGVFHRGHFKTLVLPALPLSSADKIYE
                     LHINGQSENEIVFSNRTRLTFESKSISVLIQTDKAFYKPKQEVKFRVLTLCSDLKPYR
                     TSVDIFIKDPKSNVIQQWFSQKGDLGVVSKTFQLSSNPIFGDWSIQVQVNDQQYYQSF
                     QVLEYVLPKFEVTVQTPLYCSLKSKQLNGSVIAKYTYGKPVKGSLSLTFLPLSFWGKK
                     KNITKSFEINGFANFSFDNYEMKKVMNLKPLTDVSEGSYENVDPSFPGPAEIIATVTE
                     SLTGISRMASTNVFFKQHDYIIEIFDYTTVLKPSLNFTATVKVSRSDGNQLTPEEIEN
                     DLVTVVTQRKNNHPESQRDQEMDYIQTVNYTIPQNGIIKIEFPVMSISGELQLKAYFL
                     DGTSSVTVHSMFTSPSKTYIQLKTRDEYIKVGSPFDLMVSGNRQFKDLSYMVISKGQL
                     VAAGKQSSRTFSLTPEASWAPKACIIAYYIAEDGEIINDILKIPVQLVFENKVKLFWS
                     KPTVKPSDKVSLRISATQSDSLVGIVAVDKSVTLMENSNSITMETMVHELELYNTEYY
                     LGMFMNSFAVFQECGLWVLTDATLIRDSIDEVYDTEEYSERFAEENEANLVDFEDASS
                     VNNVHVRKNFPETWIWLDAYMGSKIYEEFEVTVPDSITSWVASAFVISEDLGFGLTTV
                     PAELQAFQPFFLFLNLPYSVIRGEEFALEVSIVNYLKDTIKVVILIEESDSFDILMTS
                     NDTNGTIYRKTVQVPRDNGVTLVFPIKPTHLGEIPITVTAASPTASDAVTQTIVVKPE
                     GIEKSYSKSVLLDLTDSNVESKQQSMRFSFPPDTVIGSERVQITAIGDILGSSINGLS
                     SLIRMPYGCGEQNMIYFAPNIYILDYLTKQKQLTVNLKEKALSYMRQGYQRELLYQRE
                     DGSFSAFGDIDSSGSTWLSAFVLRCFLEADYYIDIDQDVLHRTYTWLNAHKKFNGEFW
                     EPGRVIHSELQGGTKSPVTLTAYIVTSVLGYKKYQPNIDVQDSIKFLEFEFSRGISDN
                     YTLAIISYALSTVGSPKAEEALNLLMQRSEKEGDTQFWLSSGPALSGSWQPRSVDIEI
                     AAYALLAHTLHHVSEGIPVMRWLIQQRNSLGGFVSTQDTVVALKALSEFSALVHKENT
                     DIQLTVTGPGIPRSIHFRIDSQNLFLLHQEELHALDPITVNVSAHGSGFAICQLNVDY
                     NVKGSGSSKRRRSTENQEVFDLDVIVNNEDDISHLNLNVCTSHLGSERTGMVLMEVNL
                     LSGFSASSDSIPLSETLKKVEYDNGKLNLYLDSVNESQFCVNIPTVRDYKVSNIRDGS
                     VSVMDYYEPRRQAVRSYNTQVKLSSCYLSPDTNCKSHTDGATDSLRRSSSLLVFCSVL
                     LYFVQH"
BASE COUNT     1556 a   1235 c   1251 g   1602 t
ORIGIN

  Query Match             63.2%;  Score 2709.2;  DB 10;  Length 5644;
  Best Local Similarity   77.9%;  Pred. No. 0;
  Matches 3393;  Conservative    0;  Mismatches  873;  Indels   90;  Gaps    7;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||   | |     |||||||| ||| |||||||| |||||| ||||| |  | |||||||
Db        105 ATGCGCAGTCGGAGGCTCCTGAGCGCGGCCCACCTTCTCTGCTTGTGCGCGGTCGCGCTG 164

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              |||   || || ||| | || |||||||||||||||||||||||||||| ||||||| |
Db        165 GCC---GCCCCAGGGTCCCGCTTTCTGGTGACAGCCCCAGGGATCATCAGACCCGGAGCA 221

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              |||||||||||||||||| || ||||||  | |||  |     ||||||||    ||| ||||
Db        222 AATGTGACTATTGGGGTGGATCTCCTGGAAAATAGCCCCCCACAGGTCCTTGTAAAGGCT 281

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
               || || |||||| ||| || |||      |     ||  ||||||||||||||||||||||
Db        282 CAGGTGTTCAAGATAGCTTCCAACAAATCGAGATCCATCCTGGAAGCAGAAGGAGTCTTT 341

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
                | | ||||  ||| ||||| ||     ||||||  |||||||||||| |||||||||| ||
Db        342 CACAGAGGCCATTTCAAGACTCTCGTTCTTCCGGCACTACCTCTGAGCAGTGCAGATAAG 401

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||     ||| |||||     |  || ||||||| |||||||  || || ||
Db        402 ATTTATGAGCTACACATAAACGGACAATCAGAGAATGAGATCGTATTCTCCAACAGGACA 461

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
```

```
Db      462 CGTTTAACATTTGAGAGCAAGAGTATATCTGTCCTCATTCAGACAGACAAGGCATTCTAC 521

Qy      421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
Db      522 AAGCCGAAGCAGGAAGTGAAGTTTCGGGTGCTCACACTCTGCTCAGATCTCAAGCCTTAT 581

Qy      481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
Db      582 AGGACATCCGTGGACATTTTCATTAAGGATCCGAAGTCTAATGTGATCCAACAGTGGTTT 641

Qy      541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
Db      642 TCTCAGAAAGGTGATCTTGGAGTCGTTTCCAAAACTTTTCAACTATCCTCTAATCCAATA 701

Qy      601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
Db      702 TTTGGTGACTGGTCCATTCAGGTTCAAGTGAATGATCAGCAATATTACCAGTCATTTCAG 761

Qy      661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
Db      762 GTCCTAGAATATGTGTTACCAAAGTTCGAAGTCACCGTGCAGACCCCGCTGTATTGTTCC 821

Qy      721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
Db      822 TTGAAGTCTAAACAGTTAAATGGCTCCGTCATAGCAAAGTATACATACGGGAAGCCAGTG 881

Qy      781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
Db      882 AAGGGAAGTCTGTCACTTACATTTTTACCTTTGTCTTTTGGGGAAAGAAGAAGAATATC 941

Qy      841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
Db      942 ACAAAAAGTTTGAGATAAATGGATTTGCAAACTTCTCTTTTGATAATTATGAGATGAAA 1001

Qy      901 AATGTAAT------------------GGATTCTTCAAATGGACTTTCTGAATACCTGGAT 942
Db     1002 AAGGTAATGAATTTGAAACCGCTAACGGATGTCTCGGAAGGGAGTTATGAGAATGTGGAC 1061

Qy      943 CTATCTTCCCCTGGACCAGTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATT 1002
Db     1062 CCGTCCTTTCCTGGGCCGGCAGAGATTATAGCTACTGTGACGGAATCACTCACAGGTATC 1121

Qy     1003 TCAAGAAATGTAAGCACTAATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTT 1062
Db     1122 TCAAGAATGGCAAGCACCAATGTGTTTTTCAAGCAACATGATTACATCATTGAGATTTTT 1181

Qy     1063 GATTATACTACTGTCTTGAAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGT 1122
Db     1182 GATTATACGACTGTCTTGAAACCATCTCTCAACTTCACAGCCACTGTGAAGGTCAGCCGC 1241

Qy     1123 GCTGATGGCAACCAACTGACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACA 1182
Db     1242 TCTGATGGCAATCAACTAACTCCTGAAGAAATAGAAAATGATTTGGTCACAGTAGTGACC 1301

Qy     1183 CAGAGAAACTATACTGAGTACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCT 1242
Db     1302 CAGCGGAAAAATAACCACCCTGAGAG---------TCAGAGAGATCAGGAAATGGATTAT 1352

Qy     1243 GTTCAGAAAATAAATTATACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATC 1302
Db     1353 ATCCAGACCGTAAATTATACTATCCCCCAAAATGGAATCATTAAGATTGAATTCCCAGTC 1412

Qy     1303 CTGGAGGATTCCAGTGAGCTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATG 1362
Db     1413 ATGTCCATTTCCGGCGAGCTACAACTGAAGGCTTATTTTCTTGATGGTACAAGTTCTGTG 1472

Qy     1363 GCAGTTCATAGTCTGTTTAAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGAT 1422
Db     1473 ACAGTCCACAGTATGTTCACCTCTCCTAGTAAGACGTACATCCAGCTTAAAACCAGAGAT 1532

Qy     1423 GAAAATATAAAGGTGGGATCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAG 1482
Db     1533 GAATATATAAAGGTGGGGTCACCTTTTGATTTGATGGTTAGTGGCAACAGACAATTCAAG 1592

Qy     1483 GAGTTAAGCTATATGGTAGTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCA 1542
Db     1593 GACTTAAGCTATATGGTAATATCAAAGGGACAGTTGGTGGCTGCAGGCAAACAAAGTTCA 1652

Qy     1543 ACAATGTTCTCTTTAACACCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTAT 1602
Db     1653 AGAACTTTCTCTTTAACACCAGAGGCTTCCTGGGCTCCAAAGGCCTGTATAATCGCATAT 1712

Qy     1603 TATATTGAAGATGATGGGGAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTT 1662
Db     1713 TATATTGCAGAGGATGGGGAAATTATAAATGATATTCTAAAAATCCCCGTTCAGCTTGTT 1772

Qy     1663 TTTAAAAATAAGATAAAGCTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTC 1722
Db     1773 TTTGAAAATAAGGTAAAGCTGTTTTGGAGTAAACCTACAGTCAAGCCATCTGATAAGGTC 1832

Qy     1723 TCTCTTAGGATCTCTGTGACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAA 1782
Db     1833 TCTCTCAGGATCTCTGCGACACAGTCTGACTCCCTGGTGGGGATTGTAGCTGTTGACAAA 1892

Qy     1783 AGTGTGAATCTGATGAATGCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTG 1842
Db     1893 AGTGTAACATTGATGGAAAACTCCAACAGTATTACAATGGAAACCATGGTCCATGAACTG 1952

Qy     1843 GAACTTTATAACACAGGATATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAG 1902
Db     1953 GAACTTTATAACACAGAATATTATCTAGGCATGTTCATGAACTCCTTCGCTGTATTTCAG 2012

Qy     1903 GAATGTGGACTCTGGGTATTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTT 1962
Db     2013 GAATGTGGTCTCTGGGTGTTGACAGATGCAACCCTTATAAGAGATAGCATTGATGAAGTC 2072

Qy     1963 TATGACAATGCAGAATATGCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGAT 2022
Db     2073 TACGATACTGAAGAGTATTCTGAACGCTTTGCAGAGGAAAATGAGGCAAACCTGGTTGAT 2132

Qy     2023 ATTCATGACTTTTCTTTGGGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGG 2082
Db     2133 TTTGAAGATGCTTCTTCAGTTAACAATGTACATGTCAGGAAGAATTTTCCAGAAACATGG 2192

Qy     2083 ATTTGGCTAGACACCAACATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCT 2142
Db     2193 ATTTGGCTAGACGCCTACATGGGATCTAAGATTTATGAAGAATTTGAAGTTACTGTACCT 2252

Qy     2143 GATTCTATCACTTCTTGGGTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGA 2202
Db     2253 GACTCCATCACTTCTTGGGTGGCTTCGGCTTTTGTCATCTCAGAAGACCTGGGTTTTGGG 2312

Qy     2203 CTAACAACTACTCCAGTGGAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTT 2262
Db     2313 TTAACAACCGTTCCAGCAGAGCTGCAAGCCTTCCAACCGTTTTTCCTTTTCCTGAATCTT 2372

Qy     2263 CCCTACTCTGTTATCAGAGGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTG 2322
Db     2373 CCGTACTCTGTTATCAGAGGTGAAGAGTTTGCTTTGGAAGTATCCATCGTCAATTATTTG 2432

Qy     2323 AAAGATGCCACTGAGGTTAAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATG 2382
Db     2433 AAAGATACAATTAAGGTTGTGATACTCATTGAGGAAAGTGACAGCTTTGATATTTTAATG 2492

Qy     2383 ACTTCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGAT 2442
Db     2493 ACTTCAAATGACACCAATGGCACCATATACCGGAAGACTGTTCAAGTTCCCAGGGATAAT 2552

Qy     2443 GGGGCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTC 2502
Db     2553 GGGGTCACTCTTGTTTTTCCCATCAAGCCCACGCACTTGGGAGAGATTCCCATCACAGTG 2612

Qy     2503 ACAGCTCTTTCACCCACTGCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAA 2562
Db     2613 ACGGCTGCCTCACCCACTGCCTCTGATGCTGTCACCCAGACAATTGTAGTAAAGCCTGAA 2672
```

```
Qy  2563 GGAATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAG 2622
Db  2673 GGGATAGAAAAGTCGTATTCGAAATCTGTCTTATTGGATCTGACAGATAGCAACGTAGAA 2732

Qy  2623 AGTACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGA 2682
Db  2733 AGTAAGCAGCAATCTATGAGATTCTCCTTTCCTCCTGATACGGTCATTGGCAGTGAAAGA 2792

Qy  2683 GTTCAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTG 2742
Db  2793 GTTCAGATCACAGCAATTGGAGATATCCTTGGTTCCTCCATCAATGGCTTATCTTCACTG 2852

Qy  2743 ATTCGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTAC 2802
Db  2853 ATCCGGATGCCTTATGGATGTGGTGAACAGAACATGATATATTTTGCTCCAAATATTTAC 2912

Qy  2803 ATTTTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTT 2862
Db  2913 ATTTTGGATTATCTGACTAAACAGAAACAGCTGACAGTTAATTTAAAAGAAAAGGCCCTT 2972

Qy  2863 TCATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTC 2922
Db  2973 TCATATATGAGGCAAGGTTACCAAAGGGAGCTTCTCTATCAGAGGGAAGATGGCTCCTTC 3032

Qy  2923 AGTGCTTTTGGGAATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGA 2982
Db  3033 AGTGCTTTTGGGGACATTGACTCTTCTGGGAGCACTTGGCTGTCAGCATTTGTTTTAAGA 3092

Qy  2983 TGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACACAGAACATAC 3042
Db  3093 TGCTTTCTGGAAGCTGATTACTATATAGATATTGATCAGGATGTGTTACACAGAACATAT 3152

Qy  3043 ACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATT 3102
Db  3153 ACTTGGCTTAATGCACATAAGAAATTCAATGGTGAATTTTGGGAGCCAGGAAGAGTGATT 3212

Qy  3103 CATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACT 3162
Db  3213 CACAGTGAACTTCAAGGTGGCACCAAAAGCCCAGTAACACTTACGGCCTATATTGTGACT 3272

Qy  3163 TCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTT 3222
Db  3273 TCTGTCCTGGGATACAAAAAGTATCAGCCTAATATCGATGTACAAGACTCAATCAAGTTT 3332

Qy  3223 TTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTAT 3282
Db  3333 TTGGAATTTGAATTCAGCAGAGGAATTTCGGACAATTATACCTTAGCAATTATATCCTAT 3392

Qy  3283 GCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGA 3342
Db  3393 GCCCTGTCCACAGTCGGGAGCCCTAAAGCAGAGGAGGCTTTGAACTTGCTGATGCAGCGA 3452

Qy  3343 GCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCC 3402
Db  3453 TCAGAAAAGGAAGGAGACACACAGTTCTGGTTGTCATCAGGACCTGCGCTCTCTGGGTCC 3512

Qy  3403 TGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTA 3462
Db  3513 TGGCAGCCGCGCTCCGTGGACATCGAAATCGCTGCTTATGCACTGCTGGCGCACACTCTG 3572

Qy  3463 CAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGC 3522
Db  3573 CA---CCATGTGTCTGAGGGAATCCCGGTTATGAGGTGGCTCATCCAGCAAAGAAACAGC 3629

Qy  3523 TTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTT 3582
Db  3630 CTGGGAGGTTTTGTATCCACTCAGGATACTGTTGTGGCCCTGAAGGCATTATCTGAATTT 3689

Qy  3583 GCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCA 3642
Db  3690 TCAGCCCTTGTGCACAAGGAAAATACAGATATCCAACTGACCGTCACAGGGCCCGGCATC 3749

Qy  3643 CCAAGTCCT--------------------------------------------------- 3651
Db  3750 CCCAGATCCATACACTTCCGAATTGACTCCCAGAACCTCTTTCTCCTTCACCAGGAAGAG 3809

Qy  3652 CTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATT 3711
Db  3810 CTTCATGCGCTAGATCCCATCACCGTTAATGTTTCTGCGCATGGCTCGGGATTTGCTATT 3869

Qy  3712 TGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCT 3771
Db  3870 TGCCAGCTTAATGTTGACTATAACGTGAAAGGTTCAGGTTCTTCTAAAAGACGGAGATCT 3929

Qy  3772 ATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTC 3831
Db  3930 ACCGAAAATCAAGAAGTTTTGATTTAGACGTCATTGTGA---ATAATGAGGACGACATT 3986

Qy  3832 AATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCT 3891
Db  3987 AGTCACCTGAATCTGAATGTGTGCACAAGTCACTTGGGTTCAGAGAGGACAGGCATGGTC 4046

Qy  3892 CTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGC 3951
Db  4047 CTTATGGAAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATTCAATTCCTCTGAGT 4106

Qy  3952 GAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTA 4011
Db  4107 GAGACCCTGAAGAAAGTGGAATATGATAACGGGAAACTCAACCTTTATTTAGATTCTGTA 4166

Qy  4012 AATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACC 4071
Db  4167 AATGAATCCCAGTTTTGTGTTAATATTCCCACTGTGAGAGACTACAAAGTTTCAAATATT 4226

Qy  4072 CAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGT 4131
Db  4227 CGAGATGGTTCAGTATCTGTAATGGACTACTATGAACCAAGGAGACAGGCAGTGCGAAGC 4286

Qy  4132 TACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGT 4191
Db  4287 TACAACACCCAGGTGAAGCTGTCCTCATGTTACCT---CAGTCCAGACACCAACTGCAAG 4343

Qy  4192 CCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGT 4251
Db  4344 TCACACACGGACGGAGCCACGGACTCCCTTCGACGTTCTTCGAGCCTTCTCGTCTTTTGC 4403

Qy  4252 TTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 4287
Db  4404 TCCGTCCTTCTGTACTTTGTGCAACATTGATTATGA 4439
```

```
RESULT 12
HSM805581
LOCUS       HSM805581                2938 bp    mRNA    linear   PRI 12-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp762L1111 (from clone DKFZp762L1111).
ACCESSION   AL834478
VERSION     AL834478.1  GI:21740252
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2938)
  AUTHORS   Blum,H., Bauersachs,S., Mewes,H.W., Weil,B. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by LMU (Ludwig Maximilians University,
            Munich/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            This clone (DKFZp762L1111) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
```

```
          at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .2938
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /map="15"
                     /clone="DKFZp762L1111"
                     /tissue_type="melanoma (MeWo cell line)"
                     /clone_lib="762 (synonym: hmel2). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
     gene            1. .2938
                     /gene="DKFZp762L1111"
     CDS             <1. .2565
                     /gene="DKFZp762L1111"
                     /note="similarity to alpha-2-macroglobulin (Limulus sp.)"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAD39137.1"
                     /db_xref="GI:21740253"
                     /translation="VDKSVNLMNASNDITMENVVHELELYNTGYYLGMFMNSFAVFQE
                     CGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGSSPHVRKHFPET
                     WIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPVELQAFQPFFIF
                     LNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSSEINATGHQQTLL
                     VPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAITQMILVKAEGIEKSYSQSILLD
                     LTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASLIRMPYGCGEQN
                     MINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDGSFSAFGNYDPS
                     GSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDPGRVIHSELQGG
                     NKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYTLALITYALSSV
                     GSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAAYALLSHFLQFQ
                     TSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTNIQVTVTGPSSP
                     SPVKFLIDTHNRLLLQTAELAVVQPTAVNISANGFGFAICQLNVVYNVKASGSSRRRR
                     SIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNLLSGFMVPSEAI
                     SLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDASVSIVDYYEPRR
                     QAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFKLLYFMELWL"
     polyA_signal    2843. .2848
                     /gene="DKFZp762L1111"
     polyA_site      2859
                     /gene="DKFZp762L1111"
BASE COUNT      887 a    529 c    620 g    902 t
ORIGIN

  Query Match            57.1%;  Score 2448.2;  DB 9;  Length 2938;
  Best Local Similarity  97.9%;  Pred. No. 0;
  Matches 2511;  Conservative    0;  Mismatches    3;  Indels   51;  Gaps    1;

Qy       1774 GTTGACAAAAGTGTGAATCTGATGAATGCCTCTAATGATATTACAATGGAAAATGTGGTC 1833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTTGACAAAAGTGTGAATCTGATGAATGCCTCTAATGATATTACAATGGAAAATGTGGTC 60

Qy       1834 CATGAGTTGGAACTTTATAACACAGGATATTATTTAGGCATGTTCATGAATTCTTTTGCA 1893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGAGTTGGAACTTTATAACACAGGATATTATTTAGGCATGTTCATGAATTCTTTTGCA 120

Qy       1894 GTCTTTCAGGAATGTGGACTCTGGGTATTGACAGATGCAAACCTCACGAAGGATTATATT 1953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTCTTTCAGGAATGTGGACTCTGGGTATTGACAGATGCAAACCTCACGAAGGATTATATT 180

Qy       1954 GATGGTGTTTATGACAATGCAGAATATGCTGAGAGGTTTATGGAGGAAAATGAAGGACAT 2013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GATGGTGTTTATGACAATGCAGAATATGCTGAGAGGTTTATGGAGGAAAATGAAGGACAT 240

Qy       2014 ATTGTAGATATTCATGACTTTTCTTTGGGTAGCAGTCCACATGTCCGAAAGCATTTTCCA 2073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATTGTAGATATTCATGACTTTTCTTTGGGTAGCAGTCCACATGTCCGAAAGCATTTTCCA 300

Qy       2074 GAGACTTGGATTTGGCTAGACACCAACATGGGTTACAGGATTTACCAAGAATTTGAAGTA 2133
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        301 GAGACTTGGATTTGGCTAGACACCAACATGGGTTCCAGGATTTACCAAGAATTTGAAGTA 360

Qy       2134 ACTGTACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTTGTGATCTCTGAGGACCTG 2193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 ACTGTACCTGATTCTATCACTTCTTGGGTGGCTACTGGTTTTGTGATCTCTGAGGACCTG 420

Qy       2194 GGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTCCAACCATTTTTCATTTTT 2253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAGCCTTCCAACCATTTTTCATTTTT 480

Qy       2254 TTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGAATTTGCTTTGGAAATAACTATATTC 2313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGAATCTTCCCTACTCTGTTATCAGAGGTGAAGAATTTGCTTTGGAAATAACTATATTC 540

Qy       2314 AATTATTTGAAAGATGCCACTGAGGTTAAGGTAATCATTGAGAAAAGTGACAAATTTGAT 2373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AATTATTTGAAAGATGCCACTGAGGTTAAGGTAATCATTGAGAAAAGTGACAAATTTGAT 600

Qy       2374 ATTCTAATGACTTCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCC 2433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTCTAATGACTTCAAGTGAAATAAATGCCACAGGCCACCAGCAGACCCTTCTGGTTCCC 660

Qy       2434 AGTGAGGATGGGGCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCT 2493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGTGAGGATGGGGCAACTGTTCTTTTTCCCATCAGGCCAACACATCTGGGAGAAATTCCT 720

Qy       2494 ATCACAGTCACAGCTCTTTCACCCACTGCTTCTGATGCTGTCACCCAGATGATTTTAGTA 2553
              ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db        721 ATCACAGTCACAGCTCTTTCACCCACTGCTTCTGATGCTATCACCCAGATGATTTTAGTA 780

Qy       2554 AAGGCTGAAGGAATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAAT 2613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGGCTGAAGGAATAGAAAAATCATATTCACAATCCATCTTATTAGACTTGACTGACAAT 840

Qy       2614 AGGCTACAGAGTACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGC 2673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AGGCTACAGAGTACCCTGAAAACTTTGAGTTTCTCATTTCCTCCTAATACAGTGACTGGC 900

Qy       2674 AGTGAAAGAGTTCAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTA 2733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGTGAAAGAGTTCAGATCACTGCAATTGGAGATGTTCTTGGTCCTTCCATCAATGGCTTA 960

Qy       2734 GCCTCATTGATTCGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCA 2793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCTCATTGATTCGGATGCCTTATGGCTGTGGTGAACAGAACATGATAAATTTTGCTCCA 1020

Qy       2794 AATATTTACATTTTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAA 2853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATATTTACATTTTGGATTATCTGACTAAAAAGAAACAACTGACAGATAATTTGAAAGAA 1080

Qy       2854 AAAGCTCTTTCATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGAT 2913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAAGCTCTTTCATTTATGAGGCAAGGTTACCAGAGAGAACTTCTCTATCAGAGGGAAGAT 1140

Qy       2914 GGCTCTTTCAGTGCTTTTGGGAATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTT 2973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 GGCTCTTTCAGTGCTTTTGGGAATTATGACCCTTCTGGGAGCACTTGGTTGTCAGCTTTT 1200

Qy       2974 GTTTTAAGATGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACAC 3033
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTTTTAAGATGTTTCCTTGAAGCCGATCCTTACATAGATATTGATCAGAATGTGTTACAC 1260

Qy       3034 AGAACATACACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGA 3093
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 AGAACATACACTTGGCTTAAAGGACATCAGAAATCCAACGGTGAATTTTGGGATCCAGGA 1320

Qy       3094 AGAGTGATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTAT 3153
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGAGTGATTCATAGTGAGCTTCAAGGTGGCAATAAAAGTCCAGTAACACTTACAGCCTAT 1380

Qy       3154 ATTGTAACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCT 3213
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATTGTAACTTCTCTCCTGGGATATAGAAAGTATCAGCCTAACATTGATGTGCAAGAGTCT 1440

Qy       3214 ATCCATTTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCTAGCCCTT 3273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 ATCCATTTTTTGGAGTCTGAATTCAGTAGAGGAATTTCAGACAATTATACTCTAGCCCTT 1500
```

```
Qy       3274 ATAACTTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAATATGCTG 3333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATAACTTATGCATTGTCATCAGTGGGGAGTCCTAAAGCGAAGGAAGCTTTGAATATGCTG 1560

Qy       3334 ACTTGGAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTT 3393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ACTTGGAGAGCAGAACAAGAAGGTGGCATGCAATTCTGGGTGTCATCAGAGTCCAAACTT 1620

Qy       3394 TCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCA 3453
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TCTGACTCCTGGCAGCCACGCTCCCTGGATATTGAAGTTGCAGCCTATGCACTGCTCTCA 1680

Qy       3454 CACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAA 3513
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CACTTCTTACAATTTCAGACTTCTGAGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAA 1740

Qy       3514 AGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAAGGCTCTG 3573
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 AGAAATAGCTTGGGTGGTTTTGCATCTACTCAGGATACCACTGTGGCTTTAAAGGCTCTG 1800

Qy       3574 TCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTGACGGGG 3633
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 TCTGAATTTGCAGCCCTAATGAATACAGAAAGGACAAATATCCAAGTGACCGTGACGGGG 1860

Qy       3634 CCTAGCTCACCAAGTCCT------------------------------------------ 3651
              ||||||||||||||||||
Db       1861 CCTAGCTCACCAAGTCCTGTAAAGTTTCTGATTGACACACACAACCGCTTACTCCTTCAG 1920

Qy       3652 ---------CTTGCTGTGGTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGA 3702
                       |||||||||||||||||| ||||||||||||||||||||||||||||||||
Db       1921 ACAGCAGAGCTTGCTGTGGTACAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTTGGA 1980

Qy       3703 TTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGA 3762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 TTTGCTATTTGTCAGCTCAATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGA 2040

Qy       3763 CGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAA 3822
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 CGAAGATCTATCCAAAATCAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAA 2100

Qy       3823 GATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGT 3882
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GATGATCTCAATCATGTGGATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGT 2160

Qy       3883 GGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATT 3942
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GGCATGGCTCTTATGGAAGTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATT 2220

Qy       3943 TCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTA 4002
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 TCTCTGAGCGAGACAGTGAAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTA 2280

Qy       4003 GATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTT 4062
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GATTCTGTAAATGAAACCCAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTT 2340

Qy       4063 TCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCG 4122
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 TCAAATACCCAAGATGCTTCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCG 2400

Qy       4123 GTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAG 4182
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GTGAGAAGTTACAACTCTGAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAG 2460

Qy       4183 GGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTT 4242
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GGCTGCCGTCCTTGTGAGGATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTT 2520

Qy       4243 ATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 4287
              |||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ATTTTCTGTTTCAAGCTTCTGTACTTTATGGAACTTTGGCTGTGA 2565
```

```
RESULT 13
AK095888
LOCUS       AK095888                2273 bp    mRNA    linear   PRI 15-JUL-2002
DEFINITION  Homo sapiens cDNA FLJ38569 fis, clone HCHON2006459.
ACCESSION   AK095888
VERSION     AK095888.1  GI:21755236
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
            Arita,M., Musashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
            Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
            Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
            Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B.,
            Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and
            Isogai,T.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2273)
  AUTHORS   Isogai,T. and Yamamoto,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JUL-2002) Takao Isogai, FLJ Project(HRI Team); 2-6-7
            Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
            (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            Economy, Trade and Industry of Japan; cDNA full insert sequencing:
            Research Association for Biotechnology (RAB); cDNA library
            construction: Helix Research Institute (HRI) (supported by Japan
            Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
            HRI, and Biotechnology Center, National Institute of Technology and
            Evaluation; clone selection for full insert sequencing: HRI and
            RAB; annotation: HRI and RAB.
FEATURES             Location/Qualifiers
     source          1..2273
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="HCHON2006459"
                     /cell_type="chondrocytes (HC)"
                     /clone_lib="HCHON2"
                     /note="cloning vector: pME18SFL3~primary culture,
                     chondrocytes"
     CDS             66..2063
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAC04642.1"
                     /db_xref="GI:21755237"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNLFGTQEAL"
BASE COUNT      712 a    431 c    462 g    668 t
ORIGIN

  Query Match             45.9%;  Score 1969.4;  DB 9;  Length 2273;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1970;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db         66 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 125

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 185

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        186 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 245

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 305

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 365

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 425

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 485

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 545

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 605

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 665

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        666 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAG 725

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 785

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        786 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 845

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        846 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 905

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        906 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 965

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        966 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1025

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1026 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1085

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1086 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1145

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1146 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1205

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1206 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1265

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1266 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1325

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1326 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1385

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1386 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1445

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1446 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1505

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1506 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1565

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1566 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1625

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1626 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1685

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1686 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1745

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1746 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1805

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1806 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1865

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1866 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1925

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1926 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1985

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAAT 1971
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1986 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAAT 2036
```

```
RESULT 14
AX083146
LOCUS       AX083146                1300 bp    DNA     linear   PAT 28-FEB-2001
DEFINITION  Sequence 49 from Patent WO0110903.
ACCESSION   AX083146
VERSION     AX083146.1  GI:13185055
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Yue,H., Lal,P., Tang,Y.T., Bandman,O., Baughn,M.R., Azimzai,Y.,
            Lu,D.A. and Yang,J.
  TITLE     Proteases and protease inhibitors
```

```
  JOURNAL   Patent: WO 0110903-A 49 15-FEB-2001;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .1300
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 3658034CB1"
BASE COUNT      353 a    239 c    289 g    419 t
ORIGIN

  Query Match             17.0%;  Score 730.8;  DB 6;  Length 1300;
  Best Local Similarity   93.6%;  Pred. No. 1.3e-148;
  Matches  805;  Conservative    0;  Mismatches    2;  Indels   53;  Gaps    2;

Qy       3479 AGGGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCAT 3538
              |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db          1 AGGGAATCCCAATTATGAGGCGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCAT 60

Qy       3539 CTACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATA 3598
              ||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||||
Db         61 CTACTCAGGATACCACTGTGGCTTTAAAGG--CTGTCTGAATTTGCAGCCCTAATGAATA 118

Qy       3599 CAGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        119 CAGAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGT 178

Qy       3652 --------------------------------------------CTTGCTGTGGTACAGC 3667
                                                          ||||||||||||||||
Db        179 TTCTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGC 238

Qy       3668 CAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTG 3727
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        239 CAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTG 298

Qy       3728 TATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAG 3787
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        299 TATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAG 358

Qy       3788 CCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGA 3847
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        359 CCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGA 418

Qy       3848 ATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACC 3907
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        419 ATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACC 478

Qy       3908 TATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAG 3967
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        479 TATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAG 538

Qy       3968 TGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTT 4027
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        539 TGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTT 598

Qy       4028 GTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGT 4087
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        599 GTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGT 658

Qy       4088 CCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGA 4147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        659 CCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGA 718

Qy       4148 AGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAG 4207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        719 AGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAG 778

Qy       4208 CTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACT 4267
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        779 CTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACT 838

Qy       4268 TTATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||||
Db        839 TTATGGAACTTTGGCTGTGA 858


RESULT 15
AL590428
LOCUS       AL590428               163577 bp    DNA     linear   PRI 31-JUL-2001
DEFINITION  Human DNA sequence from clone RP11-553A21 on chromosome 6, complete
            sequence.
ACCESSION   AL590428 AC026605
VERSION     AL590428.7  GI:15072593
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163577)
  AUTHORS   Chapman,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 1, 2001 this sequence version replaced gi:15021177.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr6
            RP11-553A21 is from the library RPCI-11.2 constructed by the group
            of Pieter de Jong. For further details see
            http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-553A21 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true right end of clone RP11-553A21 is at 163577 in this
            sequence. The true left end of clone RP11-525G3 is at 88067 in this
            sequence. The true right end of clone RP3-397H23 is at 2000 in this
            sequence.
FEATURES             Location/Qualifiers
     source          1. .163577
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="RP11-553A21"
                     /clone_lib="RPCI-11.2"
     repeat_region   285. .433
                     /note="AluJb repeat: matches 1. .119 of consensus"
     repeat_region   434. .743
                     /note="AluY repeat: matches 1. .311 of consensus"
     repeat_region   744. .919
                     /note="AluJb repeat: matches 119. .305 of consensus"
     repeat_region   988. .1075
                     /note="L2 repeat: matches 2626. .2708 of consensus"
     repeat_region   1166. .1384
                     /note="L1MC5 repeat: matches 7317. .7536 of consensus"
```

```
repeat_region   1385. .1690
                /note="AluSx repeat: matches 1. .306 of consensus"
repeat_region   1691. .1759
                /note="L1MC5 repeat: matches 7536. .7604 of consensus"
repeat_region   1760. .2024
                /note="AluY repeat: matches 35. .287 of consensus"
repeat_region   2025. .2248
                /note="L1MC5 repeat: matches 7604. .7887 of consensus"
repeat_region   2588. .2706
                /note="L1MC/D repeat: matches 5389. .5514 of consensus"
misc_feature    3686. .4489
                /note="CpG island"
                /evidence=not_experimental
repeat_region   3695. .3736
                /note="21 copies 2 mer cc 78% conserved"
repeat_region   4691. .4996
                /note="AluJo repeat: matches 3. .306 of consensus"
repeat_region   5628. .5977
                /note="MLT2FB repeat: matches 1. .403 of consensus"
repeat_region   5978. .6030
                /note="MER5A repeat: matches 44. .97 of consensus"
repeat_region   6118. .6474
                /note="L1MC3 repeat: matches 7398. .7739 of consensus"
repeat_region   6475. .6771
                /note="AluY repeat: matches 1. .290 of consensus"
repeat_region   6772. .7373
                /note="L1MC3 repeat: matches 6740. .7398 of consensus"
repeat_region   7491. .7764
                /note="AluJo repeat: matches 1. .286 of consensus"
repeat_region   7887. .8168
                /note="AluSq repeat: matches 1. .287 of consensus"
repeat_region   8365. .8430
                /note="33 copies 2 mer tc 89% conserved"
repeat_region   8431. .8729
                /note="AluSq repeat: matches 1. .302 of consensus"
repeat_region   8779. .8975
                /note="MER33 repeat: matches 129. .323 of consensus"
repeat_region   8976. .9292
                /note="AluSx repeat: matches 1. .311 of consensus"
repeat_region   9293. .9415
                /note="MER33 repeat: matches 1. .129 of consensus"
repeat_region   9516. .9647
                /note="MIR repeat: matches 106. .226 of consensus"
repeat_region   9698. .9994
                /note="AluSq repeat: matches 1. .295 of consensus"
repeat_region   10005. .10048
                /note="22 copies 2 mer tt 79% conserved"
repeat_region   10050. .10999
                /note="SVA repeat: matches 3. .954 of consensus"
misc_feature    10622. .11798
                /note="CpG island"
                /evidence=not_experimental
repeat_region   11056. .11951
                /note="SVA repeat: matches 519. .1386 of consensus"
repeat_region   12038. .12152
                /note="L2 repeat: matches 2588. .2707 of consensus"
repeat_region   12561. .12844
                /note="AluSx repeat: matches 3. .287 of consensus"
repeat_region   12928. .13220
                /note="AluSp repeat: matches 2. .295 of consensus"
repeat_region   13294. .13417
                /note="FLAM_C repeat: matches 1. .123 of consensus"
repeat_region   13451. .13636
                /note="AluSq repeat: matches 1. .189 of consensus"
repeat_region   13637. .13697
                /note="AluS repeat: matches 1. .63 of consensus"
repeat_region   13797. .13925
                /note="MER39 repeat: matches 15. .148 of consensus"
repeat_region   13926. .14215
                /note="AluY repeat: matches 1. .288 of consensus"
repeat_region   14216. .14662
                /note="MER39 repeat: matches 148. .543 of consensus"
repeat_region   14688. .14735
                /note="LTR29 repeat: matches 571. .619 of consensus"
repeat_region   14832. .15863
                /note="L1MA4 repeat: matches 5262. .6295 of consensus"
repeat_region   15978. .16180
                /note="L2 repeat: matches 2208. .2414 of consensus"
repeat_region   16387. .16511
                /note="L2 repeat: matches 2636. .2734 of consensus"
repeat_region   17252. .17283
                /note="16 copies 2 mer ac 96% conserved"
repeat_region   17372. .17674
                /note="AluSg repeat: matches 1. .303 of consensus"
repeat_region   18061. .19334
                /note="L1MA10 repeat: matches 5026. .6320 of consensus"
repeat_region   19352. .19437
                /note="L1M3c repeat: matches 734. .820 of consensus"
repeat_region   19545. .19620
                /note="L1M3c repeat: matches 922. .997 of consensus"
repeat_region   20319. .20442
                /note="AluJo/FRAM repeat: matches 179. .305 of consensus"
repeat_region   20921. .21277
                /note="THE1B repeat: matches 1. .364 of consensus"
repeat_region   21480. .21775
                /note="AluSg repeat: matches 1. .296 of consensus"
repeat_region   21799. .21916
                /note="FLAM_C repeat: matches 1. .118 of consensus"
repeat_region   22546. .22713
                /note="AluSg/x repeat: matches 87. .262 of consensus"
repeat_region   22895. .23001
                /note="L1MC2 repeat: matches 5827. .5924 of consensus"
repeat_region   23005. .23303
                /note="AluSx repeat: matches 14. .312 of consensus"
repeat_region   23429. .23525
                /note="AluSg/x repeat: matches 213. .309 of consensus"
repeat_region   23528. .23628
                /note="AluSq/x repeat: matches 12. .112 of consensus"
repeat_region   23626. .23695
                /note="AluSx repeat: matches 62. .142 of consensus"
repeat_region   23696. .23996
                /note="AluSg repeat: matches 1. .304 of consensus"
repeat_region   23997. .24222
                /note="AluSx repeat: matches 142. .309 of consensus"
repeat_region   24243. .24416
                /note="AluSg/x repeat: matches 127. .300 of consensus"
repeat_region   24427. .24734
                /note="AluSx repeat: matches 1. .300 of consensus"
repeat_region   24807. .25158
                /note="MLT1A1 repeat: matches 1. .363 of consensus"
repeat_region   25179. .25484
                /note="AluSx repeat: matches 1. .299 of consensus"
repeat_region   25494. .25519
                /note="13 copies 2 mer at 100% conserved"
repeat_region   25614. .25730
                /note="FLAM_C repeat: matches 2. .118 of consensus"
repeat_region   26958. .27263
                /note="AluJb repeat: matches 7. .311 of consensus"
repeat_region   27497. .27814
                /note="L1M4 repeat: matches 4499. .4833 of consensus"
repeat_region   27866. .27943
                /note="MLT1J repeat: matches 410. .492 of consensus"

  Query Match             5.4%;  Score 232;  DB 9;  Length 163577;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-40;
  Matches  232;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        277 CTACCTCTGAACAGTGCAGATGAGATTTATGAGCTACGTGTAACCGGACGTACCCAGGAT 336
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80550 CTACCTCTGAACAGTGCAGATGAGATTTATGAGCTACGTGTAACCGGACGTACCCAGGAT 80609

Qy        337 GAGATTTTATTCTCTAATAGTACCCGCTTATCATTTGAGACCAAGAGAATATCTGTCTTC 396
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80610 GAGATTTTATTCTCTAATAGTACCCGCTTATCATTTGAGACCAAGAGAATATCTGTCTTC 80669

Qy        397 ATTCAAACAGACAAGGCCTTATACAAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACA 456
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        80670 ATTCAAACAGACAAGGCCTTATACAAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACA 80729

Qy          457 CTCTTCTCAGATTTTAAGCCTTACAAAACCTCTTTAAACATTCTCATTAAGG 508
                ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        80730 CTCTTCTCAGATTTTAAGCCTTACAAAACCTCTTTAAACATTCTCATTAAGG 80781
```

Search completed: January 16, 2004, 18:48:41
Job time : 15013 secs

```
OM nucleic - nucleic search, using sw model

Run on:         January 16, 2004, 12:36:15 ; Search time 1017 Seconds
                                             (without alignments)
                                             11379.047 Million cell updates/sec

Title:          US-10-020-095-3
Perfect score:  4287
Sequence:       1 atgcagggcccaccgctcct..........ttatggaactttggctgtga 4287

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters:      5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      N_Geneseq_19Jun03:*
                1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
                25:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		4281.2	99.9	4369	25	AAD49434	Human r150 gene #1
2		4224.4	98.5	4761	24	ABQ79964	Human CD109 K1 pro
3		4224.4	98.5	5882	25	AAD49440	Human blood cell s
4		4224.4	98.5	5895	24	ABQ79966	Human CD109 K1-H7
5		4223.4	98.5	5883	25	AAD49435	Human r150 DNA #1.
6		4222.8	98.5	4761	24	ABQ79965	Human CD109 K1 var
7		4222.8	98.5	5895	24	ABQ79967	Human CD109 K1-H7
8		4221.4	98.5	4335	24	AAL49815	Human platelet all
9		4219.8	98.4	4335	24	AAL49816	Human platelet all
10		4219.6	98.4	4473	21	AAA62010	Hydrophobic domain
11		4216.6	98.4	4335	21	AAA60199	Hydrophobic domain
12		4140.2	96.6	4146	25	AAD49436	Human blood cell s
13		4082.4	95.2	4197	25	AAD49437	Human r150 DNA #2.
14		3187.4	74.4	3535	24	ABQ79968	Human CD109 K15 pr
15		3185.8	74.3	3535	24	ABQ79969	Human CD109 K15 va
16		1017	23.7	1448	24	ABL90478	Human polynucleoti
17	c	975.4	22.8	1459	22	ABA08825	Human secreted pro
18		812.2	18.9	2403	24	ABV77330	Alpha-1 proteinase
19		730.8	17.0	1300	22	AAF81735	Human protease and
20		541.6	12.6	875	20	AAX20449	Human secreted pro
21		525	12.2	821	25	ABX08843	Angiogenesis-assoc
22		497.6	11.6	744	20	AAZ15521	Human gene express
23		331.6	7.7	690	23	AAS70864	DNA encoding novel
24		300	7.0	300	20	AAZ13539	Human gene express
25		272	6.3	354	25	ABX47702	Bovine EST associa
26		175	4.1	744	20	AAZ16725	Human gene express
27		174	4.1	4527	22	AAC90951	C. elegans alpha-2
28		174	4.1	4560	22	AAC90952	C. elegans alpha-2
29		172.6	4.0	2285	23	ABL20611	Drosophila melanog
30		171.6	4.0	2480	23	ABL12153	Drosophila melanog
31		152.2	3.6	3612	23	ABL19705	Drosophila melanog
32	c	143.8	3.4	6324	23	ABL19704	Drosophila melanog
33	c	127.2	3.0	5611	23	ABL20610	Drosophila melanog
34	c	127.2	3.0	6979	23	ABL12152	Drosophila melanog
35		126.8	3.0	4615	23	AAS92193	DNA encoding novel
36		126.8	3.0	4615	24	ABN97232	Gene #3730 used to
37		126.8	3.0	4629	23	AAS92192	DNA encoding novel
38		126.8	3.0	5816	21	AAC74484	Human ORFX ORF39 p
39		125.2	2.9	4487	24	ABK92035	DNA encoding novel
40		125.2	2.9	4488	24	ABS59327	Human alpha-2-macr
41		120.8	2.8	4660	24	ABQ93897	Human pregnancy zo
42		120.6	2.8	2608	24	AAL49817	Human platelet all
43		119.8	2.8	4501	25	ABX70464	DNA encoding human
44		119.4	2.8	1140	22	ABA06696	Human cDNA SEQ ID
45		119.4	2.8	1140	22	ABA06869	Human polynucleoti

ALIGNMENTS

```
RESULT 1
AAD49434
ID   AAD49434 standard; DNA; 4369 BP.
XX
AC   AAD49434;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 gene #1.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             26..4312
FT                   /*tag=  a
FT                   /product= "Human r1520 protein #1"
FT                   /transl_except= (pos:2132..2134, aa:Xaa)
FT                   /note= "Xaa corresponds to Ser, Tyr"
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
```

```
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   P-PSDB; AAE32012.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 91-93; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is a gene encoding
CC   human r150 protein.
XX
SQ   Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;

  Query Match             99.9%;  Score 4281.2;  DB 25;  Length 4369;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 4283;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 85

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 145

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 205

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 265

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 325

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 385

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 445

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 505

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 565

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 625

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATANTATCAATCATTTCAG 660
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        626 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAG 685

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 745

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 805

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 865

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 925

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 985

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1045

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1105

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1165

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1225

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1285

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1345

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1405

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1406 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1465

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1466 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1525

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1526 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1585

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1586 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1645

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1646 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1705

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1706 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1765
```

```
Qy   1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1766 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1825

Qy   1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1826 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1885

Qy   1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1886 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1945

Qy   1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1946 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2005

Qy   1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2006 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2065

Qy   2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2066 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2125

Qy   2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2126 ATGGGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2185

Qy   2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2186 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2245

Qy   2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2246 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2305

Qy   2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2306 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2365

Qy   2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2366 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2425

Qy   2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2426 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2485

Qy   2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2486 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2545

Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2546 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2605

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2606 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2665

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2666 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2725

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2726 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2785

Qy   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTGGATTATCTGACT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2786 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTGGATTATCTGACT 2845

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2846 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2905

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2906 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2965

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2966 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3025

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3026 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3085

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3086 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3145

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3146 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3205

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3206 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3265

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3266 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3325

Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3326 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3385

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3386 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3445

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3446 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3505

Qy   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3506 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3565

Qy   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3566 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3625

Qy   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTG 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3626 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTG 3685

Qy   3661 GTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTC 3720
          |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Db   3686 GTACAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTC 3745

Qy   3721 AATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAAT 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3746 AATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAAT 3805

Qy   3781 CAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTG 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3806 CAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTG 3865

Qy   3841 GATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAA 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3866 GATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAA 3925

Qy   3901 GTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTG 3960
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3926 GTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTG 3985

Qy     3961 AAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACC 4020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3986 AAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACC 4045

Qy     4021 CAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCT 4080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4046 CAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCT 4105

Qy     4081 TCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCT 4140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4106 TCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCT 4165

Qy     4141 GAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAG 4200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4166 GAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAG 4225

Qy     4201 GATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTT 4260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4226 GATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTT 4285

Qy     4261 CTGTACTTTATGGAACTTTGGCTGTGA 4287
            |||||||||||||||||||||||||||
Db     4286 CTGTACTTTATGGAACTTTGGCTGTGA 4312
```

```
RESULT 2
ABQ79964
ID   ABQ79964 standard; cDNA; 4761 BP.
XX
AC   ABQ79964;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1 protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82165.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 1a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1 cDNA sequence.
XX
SQ   Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

  Query Match             98.5%;  Score 4224.4;  DB 24;  Length 4761;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy        1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy       61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy      121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy      181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy      241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy      301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy      361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy      421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy      481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy      541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy      601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy      661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy    721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy    781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy    841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy    901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy    961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy   1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1192

Qy   1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy   1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy   1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy   1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy   1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy   1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy   1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy   1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy   1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy   1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy   1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy   1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy   1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy   1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy   1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy   1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy   2041 GGTAGCAGTCCACATGTCCGAAAGCATTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2153 GGTAGCAGTCCACATGTCCGAAAGCATTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy   2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy   2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy   2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy   2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy   2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy   2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy   2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992
```

```
Qy       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 3
AAD49440
ID   AAD49440 standard; DNA; 5882 BP.
XX
AC   AAD49440;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human blood cell surface antigen, CD109 encoding DNA #2.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; blood cell syrface antigen; CD109; ds.
XX
OS   Homo sapiens.
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Disclosure; Fig 17; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
```

```
CC   and increase graft success. The present sequence is human blood cell
CC   surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.
CC   Note: This sequence is stated to encode human CD109 protein shown in
CC   figure 17 of the specification, but this does not appear to be the case.
XX
SQ   Sequence 5882 BP; 1696 A; 1173 C; 1241 G; 1772 T; 0 other;

  Query Match             98.5%;  Score 4224.4;  DB 25;  Length 5882;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152
```

```
Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy         4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy         4270 ATGGAACTTTGGCTGTGA 4287
                ||||||||||||||||||
Db         4433 ATGGAACTTTGGCTGTGA 4450


RESULT 4
ABQ79966
ID   ABQ79966 standard; cDNA; 5895 BP.
XX
AC   ABQ79966;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1-H7 protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1-H7; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1-H7"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82167.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 2a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1-H7 cDNA sequence.
XX
SQ   Sequence 5895 BP; 1709 A; 1173 C; 1241 G; 1772 T; 0 other;

  Query Match             98.5%;  Score 4224.4;  DB 24;  Length 5895;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy            1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy           61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy          121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy          181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy          241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy          301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy          361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy          421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy          481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy          541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy          601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy          661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy          721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy          781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy          841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy          901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy      961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy     1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1192

Qy     1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy     1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy     1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy     1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy     1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy     1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy     1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy     1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy     1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy     1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy     1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy     1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy     1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy     1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy     1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy     1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy     2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy     2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy     2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy     2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy     2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy     2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy     2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy     2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy     2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy     2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy     2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy     2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy     2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy     2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy     2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy     2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy     3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy     3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232
```

```
Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 5
AAD49435
ID   AAD49435 standard; DNA; 5883 BP.
XX
AC   AAD49435;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 DNA #1.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "Human r1520 protein #2"
FT                   /transl_except= (pos:2219..2221, aa:Xaa)
FT                   /note= "Xaa corresponds to Ser, Tyr"
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   P-PSDB; AAE32013.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 100-103; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human r150 protein
CC   encoding DNA.
XX
SQ   Sequence 5883 BP; 1696 A; 1173 C; 1241 G; 1772 T; 1 other;

  Query Match             98.5%;  Score 4223.4;  DB 25;  Length 5883;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4285;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;
```

```
Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 6
ABQ79965
ID   ABQ79965 standard; cDNA; 4761 BP.
XX
AC   ABQ79965;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1 variant protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1; variant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1 variant"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82166.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 1b; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1 variant cDNA sequence.
XX
SQ   Sequence 4761 BP; 1419 A; 913 C; 1016 G; 1413 T; 0 other;

  Query Match             98.5%;  Score 4222.8;  DB 24;  Length 4761;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4285;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1192
```

```
Qy   1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy   1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy   1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy   1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy   1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy   1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy   1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy   1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy   1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy   1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy   1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy   1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy   1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy   1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy   1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy   1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy   2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy   2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy   2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy   2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy   2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy   2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy   2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy   2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                        ||||||||||||||||||
Db       3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 7
ABQ79967
ID   ABQ79967 standard; cDNA; 5895 BP.
XX
AC   ABQ79967;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1-H7 variant protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1-H7; variant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1-H7 variant"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82168.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 2b; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1-H7 variant cDNA sequence.
XX
SQ   Sequence 5895 BP; 1708 A; 1174 C; 1241 G; 1772 T; 0 other;

  Query Match             98.5%;  Score 4222.8;  DB 24;  Length 5895;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4285;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;
```

```
Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy        2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy        2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy        2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
               ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy        2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy        2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy        2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy        2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy        2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy        2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy        2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy        2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy        2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy        2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy        3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy        3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy        3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy        3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy        3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy        3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy        3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy        3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy        3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3593 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy        3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy        3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy        3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                         ||||||||||||||||||
Db        3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy        3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy        3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy        3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy        3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy        3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy        3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy        4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy        4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy        4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy        4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy        4270 ATGGAACTTTGGCTGTGA 4287
               ||||||||||||||||||
Db        4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 8
AAL49815
ID   AAL49815 standard; cDNA; 4335 BP.
XX
AC   AAL49815;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human platelet alloantigen Gova coding sequence.
XX
KW   Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;
KW   SNP; biallelic; bleeding disorder; post-transfusion purpura;
KW   post-transfusion platelet refractoriness; haemostatic; vaccine;
KW   neonatal alloimmune thrombocytopenia; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4335
FT                   /*tag=  a
FT                   /product= "Gova"
FT                   /partial
FT                   /note= "no stop codon"
FT   variation       replace(2108,C)
FT                   /*tag=  b
XX
PN   WO200270738-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00291.
XX
PR   07-MAR-2001; 2001US-273941P.
XX
PA   (SCHU/) SCHUH A.
XX
PI   Schuh A,  Ouwehand W;
XX
DR   WPI; 2002-713460/77.
DR   P-PSDB; AAO19372.
XX
PT   New isolated oligonucleotide binding to a region of CD109 nucleic acid
PT   having a single nucleotide polymorphism that distinguishes a Gova
PT   and/or Govb allele, useful for treating blood disorders e.g. alloimmune
PT   thrombocytopenia  -
XX
PS   Claim 6; Page 23-29; 69pp; English.
XX
CC   The present invention relates to a sequence capable off binding
CC   specifically to a CD109 nucleic acid which has a single nucleotide
CC   polymorphism that distinguishes the Gova and Govb alleles. Detection of
CC   the Gov genotype is useful for detecting whether the subject has or is at
CC   risk of a blood disease, disorder or abnormal physical state, such as
CC   bleeding, or increased risk of bleeding, due to alloimmune destruction of
CC   blood platelets, e.g., post-transfusion purpura, post-transfusion
CC   platelet refractoriness or neonatal alloimmune thrombocytopenia. The
CC   nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
CC   individuals. The present sequence is the Gova coding sequence.
XX
SQ   Sequence 4335 BP; 1296 A; 837 C; 922 G; 1280 T; 0 other;

  Query Match             98.5%;  Score 4221.4;  DB 24;  Length 4335;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4283;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
```

```
Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400

Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300

Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy        3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy        3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy        3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy        3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy        3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                          ||||||||||||||||||
Db        3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy        3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy        3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy        3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy        3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy        3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy        3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy        4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy        4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy        4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy        4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy        4270 ATGGAACTTTGGCTG 4284
               |||||||||||||||
Db        4321 ATGGAACTTTGGCTG 4335


RESULT 9
AAL49816
ID   AAL49816 standard; cDNA; 4335 BP.
XX
AC   AAL49816;
XX
DT   04-DEC-2002  (first entry)
XX
DE   Human platelet alloantigen Govb coding sequence.
XX
KW   Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;
KW   SNP; biallelic; bleeding disorder; post-transfusion purpura;
KW   post-transfusion platelet refractoriness; haemostatic; vaccine;
KW   neonatal alloimmune thrombocytopenia; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4335
FT                   /*tag=  a
FT                   /product= "Govb"
FT                   /partial
FT                   /note= "no stop codon"
FT   variation       replace(2108,A)
FT                   /*tag=  b
XX
PN   WO200270738-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00291.
XX
PR   07-MAR-2001; 2001US-273941P.
XX
PA   (SCHU/) SCHUH A.
XX
PI   Schuh A,  Ouwehand W;
XX
DR   WPI; 2002-713460/77.
DR   P-PSDB; AAO19373.
XX
PT   New isolated oligonucleotide binding to a region of CD109 nucleic acid
PT   having a single nucleotide polymorphism that distinguishes a Gova
PT   and/or Govb allele, useful for treating blood disorders e.g. alloimmune
PT   thrombocytopenia  -
XX
PS   Claim 7; Page 35-41; 69pp; English.
XX
CC   The present invention relates to a sequence capable off binding
CC   specifically to a CD109 nucleic acid which has a single nucleotide
CC   polymorphism that distinguishes the Gova and Govb alleles. Detection of
CC   the Gov genotype is useful for detecting whether the subject has or is at
CC   risk of a blood disease, disorder or abnormal physical state, such as
CC   bleeding, or increased risk of bleeding, due to alloimmune destruction of
CC   blood platelets, e.g., post-transfusion purpura, post-transfusion
CC   platelet refractoriness or neonatal alloimmune thrombocytopenia. The
CC   nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
CC   individuals. The present sequence is the Govb coding sequence.
XX
SQ   Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other;

  Query Match             98.4%;  Score 4219.8;  DB 24;  Length 4335;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4282;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;

Qy           1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy          61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy         121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy         181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTAAAAATAAGATAAAG 1680

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400
```

```
Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300

Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                          ||||||||||||||||||
Db       3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy       4270 ATGGAACTTTGGCTG 4284
              |||||||||||||||
Db       4321 ATGGAACTTTGGCTG 4335
```

```
RESULT 10
AAA62010
ID   AAA62010 standard; DNA; 4473 BP.
XX
AC   AAA62010;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.
XX
KW   Human; secreted protein; membrane protein; hydrophobic domain;
KW   proliferation control; differentiation induction; material transport;
```

```
KW   biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
KW   immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
KW   haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
KW   autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200029448-A2.
XX
PD   25-MAY-2000.
XX
PF   17-NOV-1999;   99WO-JP06412.
XX
PR   17-NOV-1998;   98JP-0326255.
PR   22-DEC-1998;   98JP-0364315.
PR   16-MAR-1999;   99JP-0069811.
PR   27-APR-1999;   99JP-0119299.
PR   19-MAY-1999;   99JP-0138169.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Kimura T;
XX
DR   WPI; 2000-387753/33.
DR   P-PSDB; AAB12127.
XX
PT   Proteins comprising hydrophobic regions, such as secretory and membrane
PT   proteins, useful in research and diagnostics and having various
PT   activities e.g. immunomodulatory, antiinflammatory, chemokinetic,
PT   hemostatic, thrombolytic -
XX
PS   Claim 4; Page 253-261; 410pp; English.
XX
CC   Secretory proteins play important roles in the proliferation control, the
CC   differentiation induction, the material transport and the biophylaxis of
CC   cells. Membrane proteins have important roles as signal receptors, ion
CC   channels and transporters. The present sequence is the coding sequence
CC   for a human protein which has at least one hydrophobic domain. The
CC   protein encoded by the present sequence may be a secretory or a membrane
CC   protein. The encoded protein may have cytokine and cell
CC   proliferation/differentiation activity, immune stimulating or suppressing
CC   activity, haematopoiesis activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, anti-inflammatory activity and tumour
CC   inhibition activity. The present sequence could therefore be used for
CC   treatment of autoimmune disease, Alzheimer's disease, Parkinson's
CC   disease, and cancer via gene therapy.
XX
SQ   Sequence 4473 BP; 1328 A; 867 C; 952 G; 1326 T; 0 other;

  Query Match             98.4%;  Score 4219.6;  DB 21;  Length 4473;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 4283;  Conservative    0;  Mismatches    4;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         45 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 104

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        105 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 164

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        165 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 224

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        225 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 284

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        285 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 344

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        345 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 404

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        405 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 464

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        465 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 524

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        525 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 584

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        585 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 644

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        645 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 704

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        705 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 764

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        765 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 824

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        825 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 884

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        885 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 944

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db        945 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTTCCCTGGACCA 1004

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1005 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1064

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1065 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1124

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1125 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1184

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1185 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1244

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1245 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1304

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1305 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1364

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1365 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1424

Qy   1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1425 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1484

Qy   1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1485 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1544

Qy   1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1545 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1604

Qy   1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1605 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1664

Qy   1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1665 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1724

Qy   1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1725 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1784

Qy   1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1785 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1844

Qy   1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1845 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1904

Qy   1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1905 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1964

Qy   1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1965 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2024

Qy   1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2025 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2084

Qy   2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2085 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2144

Qy   2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2145 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2204

Qy   2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2205 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2264

Qy   2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2265 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2324

Qy   2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2325 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2384

Qy   2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2385 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2444

Qy   2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2445 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2504

Qy   2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2505 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2564

Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   2565 GCTTCTGATGCTATCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2624

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2625 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2684

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2685 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2744

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2745 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2804

Qy   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2805 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2864

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2865 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2924

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2925 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2984

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2985 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3044

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3045 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3104

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3105 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3164

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3165 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3224

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3225 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3284

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3285 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3344

Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3345 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3404

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3405 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3464

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3465 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3524

Qy   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3525 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3584
```

```
Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3585 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3644

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3645 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3704

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                        ||||||||||||||||||
Db       3705 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3764

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3765 ACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3824

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3825 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3884

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3885 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3944

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3945 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4004

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4005 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4064

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4065 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4124

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4125 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4184

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4185 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4244

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4245 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4304

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4269
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4305 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4364

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4365 ATGGAACTTTGGCTGTGA 4382
```

```
RESULT 11
AAA60199
ID   AAA60199 standard; DNA; 4335 BP.
XX
AC   AAA60199;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.
XX
KW   Human; secreted protein; membrane protein; hydrophobic domain;
KW   proliferation control; differentiation induction; material transport;
KW   biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
KW   immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
KW   haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
KW   autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200029448-A2.
XX
PD   25-MAY-2000.
XX
PF   17-NOV-1999;  99WO-JP06412.
XX
PR   17-NOV-1998;  98JP-0326255.
PR   22-DEC-1998;  98JP-0364315.
PR   16-MAR-1999;  99JP-0069811.
PR   27-APR-1999;  99JP-0119299.
PR   19-MAY-1999;  99JP-0138169.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S,  Kimura T;
XX
DR   WPI; 2000-387753/33.
DR   P-PSDB; AAB12127.
XX
PT   Proteins comprising hydrophobic regions, such as secretory and membrane
PT   proteins, useful in research and diagnostics and having various
PT   activities e.g. immunomodulatory, antiinflammatory, chemokinetic,
PT   hemostatic, thrombolytic -
XX
PS   Claim 3; Page 244-246; 410pp; English.
XX
CC   Secretory proteins play important roles in the proliferation control, the
CC   differentiation induction, the material transport and the biophylaxis of
CC   cells. Membrane proteins have important roles as signal receptors, ion
CC   channels and transporters. The present sequence is the coding sequence
CC   for a human protein which has at least one hydrophobic domain. The
CC   protein encoded by the present sequence may be a secretory or a membrane
CC   protein. The encoded protein may have cytokine and cell
CC   proliferation/differentiation activity, immune stimulating or suppressing
CC   activity, haematopoiesis activity, tissue growth activity,
CC   activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, anti-inflammatory activity and tumour
CC   inhibition activity. The present sequence could therefore be used for
CC   treatment of autoimmune disease, Alzheimer's disease, Parkinson's
CC   disease, and cancer via gene therapy.
XX
SQ   Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other;

  Query Match             98.4%;  Score 4216.6;  DB 21;  Length 4335;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 4280;  Conservative    0;  Mismatches    4;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy         361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy         421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy         481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy         541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy         601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy         661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy         721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy         781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840

Qy         841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy         901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
                 |||||||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db         901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTTCCCTGGACCA 960

Qy         961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy        1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy        1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy        1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy        1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy        1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy        1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy        1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy        1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy        1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy        1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy        1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy        1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy        1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy        1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy        1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy        1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy        1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy        2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy        2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
                 ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2101 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy        2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy        2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy        2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy        2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400

Qy        2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy        2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
```

```
Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
          |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 GCTTCTGATGCTATCACCCAGATGATTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300

Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy   3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                    ||||||||||||||||||
Db   3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy   3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 ACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy   3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy   3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy   3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy   3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy   3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy   4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy   4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy   4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy   4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy   4270 ATGGAACTTTGGCTG 4284
          |||||||||||||||
Db   4321 ATGGAACTTTGGCTG 4335
```

```
RESULT 12
AAD49436
ID   AAD49436 standard; DNA; 4146 BP.
XX
AC   AAD49436;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human blood cell surface antigen, CD109 encoding DNA #1.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; blood cell syrface antigen; CD109; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4146
FT                   /*tag=  a
FT                   /product= "Human CD109 protein"
```

```
FT                   /transl_except= (pos:2044..2046, aa:Xaa)
FT                   /note= "Xaa corresponds to Ser, Tyr; No start
FT                   and stop codon"
FT                   /partial
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   P-PSDB; AAE32014.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 109-112; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human blood cell
CC   surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.
XX
SQ   Sequence 4146 BP; 1260 A; 776 C; 885 G; 1224 T; 1 other;

  Query Match             96.6%;  Score 4140.2;  DB 25;  Length 4146;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 4142;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         64 GTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGAAAT 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGAAAT 60

Qy        124 GTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCGGAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCGGAG 120

Qy        184 CTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTTGAA 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTTGAA 180

Qy        244 AAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAGATT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAGATT 240

Qy        304 TATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACCCGC 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACCCGC 300

Qy        364 TTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATACAAG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATACAAG 360

Qy        424 CCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTACAAA 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTACAAA 420

Qy        484 ACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTGTCA 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTGTCA 480

Qy        544 CAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATACTT 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATACTT 540

Qy        604 GGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAGGTT 663
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db        541 GGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAGGTT 600

Qy        664 TCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCTATG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCTATG 660

Qy        724 AATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTGAAA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTGAAA 720

Qy        784 GGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATTACA 843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATTACA 780

Qy        844 AAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAAAT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAAAT 840

Qy        904 GTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCAGTA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCAGTA 900

Qy        964 GAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACTAAT 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACTAAT 960

Qy       1024 GTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTGAAG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTGAAG 1020

Qy       1084 CCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTGACT 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTGACT 1080

Qy       1144 CTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAGTAC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAGTAC 1140

Qy       1204 TGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTATACT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTATACT 1200

Qy       1264 GTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAGCTA 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAGCTA 1260

Qy       1324 CAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTTAAG 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTTAAG 1320

Qy       1384 TCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGATCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGATCG 1380

Qy       1444 CCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTAGTA 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTAGTA 1440

Qy       1504 TCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACACCA 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACACCA 1500

Qy       1564 GAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGGGAA 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     1501 GAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGGGAA 1560

Qy     1624 ATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAGCTA 1683
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 ATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAGCTA 1620

Qy     1684 TATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTGACA 1743
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 TATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTGACA 1680

Qy     1744 CAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAATGCC 1803
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 CAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAATGCC 1740

Qy     1804 TCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGATAT 1863
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 TCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGATAT 1800

Qy     1864 TATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTATTG 1923
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 TATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTATTG 1860

Qy     1924 ACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATATGCT 1983
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 ACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATATGCT 1920

Qy     1984 GAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTGGGT 2043
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 GAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTGGGT 1980

Qy     2044 AGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATG 2103
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 AGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATG 2040

Qy     2104 GGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTG 2163
            |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 GGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTG 2100

Qy     2164 GCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAG 2223
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAG 2160

Qy     2224 CTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGT 2283
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 CTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGT 2220

Qy     2284 GAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAG 2343
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 GAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAG 2280

Qy     2344 GTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAATGCC 2403
            |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db     2281 GTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAATGCC 2340

Qy     2404 ACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTTCCC 2463
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2341 ACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTTCCC 2400

Qy     2464 ATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACTGCT 2523
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 ATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACTGCT 2460

Qy     2524 TCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATATTCA 2583
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 TCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATATTCA 2520

Qy     2584 CAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGT 2643
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 CAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGT 2580

Qy     2644 TTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGA 2703
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 TTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGA 2640

Qy     2704 GATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGT 2763
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 GATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGT 2700

Qy     2764 GGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAA 2823
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 GGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAA 2760

Qy     2824 AAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTAC 2883
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 AAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTAC 2820

Qy     2884 CAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGAC 2943
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 CAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGAC 2880

Qy     2944 CCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCT 3003
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 CCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCT 2940

Qy     3004 TACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAG 3063
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 TACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAG 3000

Qy     3064 AAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGC 3123
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 AAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGC 3060

Qy     3124 AATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAG 3183
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 AATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAG 3120

Qy     3184 TATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGA 3243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 TATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGA 3180

Qy     3244 GGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGT 3303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 GGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGT 3240

Qy     3304 CCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATG 3363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3241 CCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATG 3300

Qy     3364 CAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGAT 3423
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3301 CAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGAT 3360

Qy     3424 ATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGA 3483
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3361 ATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGA 3420

Qy     3484 ATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACT 3543
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3421 ATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACT 3480

Qy     3544 CAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAA 3603
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3481 CAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAA 3540

Qy     3604 AGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTGGTA 3663
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3541 AGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTGGTA 3600

Qy     3664 CAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAAT 3723
            ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3601 CAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAAT 3660

Qy     3724 GTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAA 3783
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3661 GTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAA 3720
```

```
Qy       3784 GAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGAT 3843
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 GAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGAT 3780

Qy       3844 TTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTT 3903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTT 3840

Qy       3904 AACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAG 3963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 AACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAG 3900

Qy       3964 AAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAG 4023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 AAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAG 3960

Qy       4024 TTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCA 4083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCA 4020

Qy       4084 GTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAA 4143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 GTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAA 4080

Qy       4144 GTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGAT 4203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGAT 4140

Qy       4204 GGAGCT 4209
              ||||||
Db       4141 GGAGCT 4146

RESULT 13
AAD49437
ID   AAD49437 standard; DNA; 4197 BP.
XX
AC   AAD49437;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 DNA #2.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4197
FT                   /*tag=  a
FT                   /product= "Human r1520 protein #3"
FT                   /transl_except= (pos:2044..2046, aa:Xaa)
FT                   /note= "Xaa corresponds to Ser, Tyr; No start
FT                   and stop codon"
FT                   /partial
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   P-PSDB; AAE32015.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 118-120; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is human r150 protein
CC   encoding DNA.
XX
SQ   Sequence 4197 BP; 1276 A; 788 C; 894 G; 1238 T; 1 other;

  Query Match             95.2%;  Score 4082.4;  DB 25;  Length 4197;
  Best Local Similarity   98.7%;  Pred. No. 0;
  Matches 4144;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;

Qy         64 GTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGAAAT 123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGAAAT 60

Qy        124 GTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCGGAG 183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCGGAG 120

Qy        184 CTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTTGAA 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTTGAA 180

Qy        244 AAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAGATT 303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAGATT 240

Qy        304 TATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACCCGC 363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACCCGC 300

Qy        364 TTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATACAAG 423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATACAAG 360

Qy        424 CCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTACAAA 483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTACAAA 420

Qy        484 ACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTGTCA 543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTGTCA 480

Qy        544 CAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATACTT 603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATACTT 540

Qy        604 GGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAGGTT 663
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAGGTT 600

Qy        664 TCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCTATG 723
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCTATG 660

Qy        724 AATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTGAAA 783
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTGAAA 720

Qy        784 GGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATTACA 843
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GGAGACGTAACGCTTACATTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATTACA 780

Qy        844 AAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAAAT 903
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAAAAT 840

Qy        904 GTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCAGTA 963
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCAGTA 900

Qy        964 GAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACTAAT 1023
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACTAAT 960

Qy       1024 GTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTGAAG 1083
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTGAAG 1020

Qy       1084 CCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTGACT 1143
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTGACT 1080

Qy       1144 CTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAGTAC 1203
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAGTAC 1140

Qy       1204 TGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTATACT 1263
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTATACT 1200

Qy       1264 GTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAGCTA 1323
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 GTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAGCTA 1260

Qy       1324 CAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTTAAG 1383
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTTAAG 1320

Qy       1384 TCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGATCG 1443
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 TCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGATCG 1380

Qy       1444 CCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTAGTA 1503
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTAGTA 1440

Qy       1504 TCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACACCA 1563
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACACCA 1500

Qy       1564 GAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGGGAA 1623
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGGGAA 1560

Qy       1624 ATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAGCTA 1683
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAGCTA 1620

Qy       1684 TATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTGACA 1743
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 TATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTGACA 1680

Qy       1744 CAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAATGCC 1803
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAATGCC 1740

Qy       1804 TCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGATAT 1863
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGATAT 1800

Qy       1864 TATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTATTG 1923
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db       1801 TATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTATTG 1860

Qy       1924 ACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATATGCT 1983
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATATGCT 1920

Qy       1984 GAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTGGGT 2043
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTGGGT 1980

Qy       2044 AGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATG 2103
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATG 2040

Qy       2104 GGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTG 2163
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGGGTG 2100

Qy       2164 GCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAG 2223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 GCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTGGAG 2160

Qy       2224 CTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGT 2283
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 CTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGAGGT 2220

Qy       2284 GAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAG 2343
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTTAAG 2280

Qy       2344 GTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAATGCC 2403
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db       2281 GTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAATGCC 2340

Qy       2404 ACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTTCCC 2463
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 ACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTTCCC 2400

Qy       2464 ATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACTGCT 2523
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 ATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACTGCT 2460

Qy       2524 TCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATATTCA 2583
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 TCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATATTCA 2520

Qy       2584 CAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGT 2643
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 CAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTGAGT 2580

Qy       2644 TTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGA 2703
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 TTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATTGGA 2640

Qy       2704 GATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGT 2763
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGCTGT 2700

Qy       2764 GGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAA 2823
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 GGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACTAAA 2760

Qy       2824 AAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTAC 2883
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 AAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGTTAC 2820

Qy       2884 CAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGAC 2943
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 CAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTATGAC 2880

Qy       2944 CCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCT 3003
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGATCCT 2940
```

```
Qy       3004 TACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAG 3063
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 TACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACATCAG 3000

Qy       3064 AAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGC 3123
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 AAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGTGGC 3060

Qy       3124 AATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAG 3183
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 AATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGAAAG 3120

Qy       3184 TATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGA 3243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 TATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGTAGA 3180

Qy       3244 GGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGT 3303
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3181 GGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGGAGT 3240

Qy       3304 CCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATG 3363
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3241 CCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGCATG 3300

Qy       3364 CAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGAT 3423
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 CAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTGGAT 3360

Qy       3424 ATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGA 3483
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAGGGA 3420

Qy       3484 ATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACT 3543
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 ATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCTACT 3480

Qy       3544 CAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAA 3603
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 CAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACAGAA 3540

Qy       3604 AGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT------------ 3651
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 AGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTTCTG 3600

Qy       3652 ---------------------------------------CTTGCTGTGGTACAGCCAATG 3672
                                                     |||||||||||||||||||||
Db       3601 ATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCAATG 3660

Qy       3673 GCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATAT 3732
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3661 GCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTATAT 3720

Qy       3733 AATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTT 3792
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 AATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCCTTT 3780

Qy       3793 GATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTG 3852
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 GATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAATGTG 3840

Qy       3853 TGTACAAGCTTTTCGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTA 3912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 TGTACAAGCTTTTCGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTATTA 3900

Qy       3913 AGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAA 3972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 AGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGGAA 3960

Qy       3973 TATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTT 4032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGTGTT 4020

Qy       4033 AATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATA 4092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 AATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCCATA 4080

Qy       4093 GTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTG 4152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAGCTG 4140

Qy       4153 TCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 TCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4197
```

```
RESULT 14
ABQ79968
ID   ABQ79968 standard; cDNA; 3535 BP.
XX
AC   ABQ79968;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K15 protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K15; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K15"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82169.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 4a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
```

```
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K15 cDNA sequence.
XX
SQ   Sequence 3535 BP; 1108 A; 675 C; 731 G; 1021 T; 0 other;

  Query Match             74.4%;  Score 3187.4;  DB 24;  Length 3535;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 3188;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy     2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy     2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy     2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy     2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy     2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy     2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy     2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy     2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy     2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy     2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy     2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy     2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy     2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy     2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy     2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy     2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy     3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy     3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy     3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy     3181 AAGTATCAG 3189
            |||||||||
Db     3293 AAGTATCAG 3301
```

```
RESULT 15
ABQ79969
ID   ABQ79969 standard; cDNA; 3535 BP.
XX
AC   ABQ79969;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K15 variant protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K15; variant; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K15 variant"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82170.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 4b; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
```

```
CC   treating these conditions. The present sequence represents the human
CC   CD109 K15 variant cDNA sequence.
XX
SQ   Sequence 3535 BP; 1107 A; 676 C; 731 G; 1021 T; 0 other;

  Query Match             74.3%;  Score 3185.8;  DB 24;  Length 3535;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 3187;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy         2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
                ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2213 ATGGGTTCCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy         2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy         2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy         2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy         2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
                ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db         2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy         2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy         2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632

Qy         2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy         2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy         2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy         2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy         2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy         2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy         2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy         2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy         3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy         3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy         3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy         3181 AAGTATCAG 3189
                |||||||||
Db         3293 AAGTATCAG 3301
```

Search completed: January 16, 2004, 14:38:44
Job time : 1046 secs

```
OM nucleic - nucleic search, using sw model

Run on:         January 16, 2004, 12:36:15 ; Search time 1017 Seconds
                                             (without alignments)
                                             11379.047 Million cell updates/sec

Title:          US-10-020-095-3
Perfect score:  4287
Sequence:       1 atgcagggcccaccgctcct..........ttatggaactttggctgtga 4287

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters:      5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_19Jun03:*
              1:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
              2:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
              3:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
              4:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
              5:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
              6:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
              7:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
              8:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
              9:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
             10:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
             11:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
             12:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
             13:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
             14:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
             15:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
             16:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
             17:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
             18:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
             19:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
             20:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
             21:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
             22:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
             23:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
             24:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
             25:   /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	4281.2	99.9	4369	25	AAD49434	Human r150 gene #1
	2	4224.4	98.5	4761	24	ABQ79964	Human CD109 K1 pro
	3	4224.4	98.5	5882	25	AAD49440	Human blood cell s
	4	4224.4	98.5	5895	24	ABQ79966	Human CD109 K1-H7
	5	4223.4	98.5	5883	25	AAD49435	Human r150 DNA #1.
	6	4222.8	98.5	4761	24	ABQ79965	Human CD109 K1 var
	7	4222.8	98.5	5895	24	ABQ79967	Human CD109 K1-H7
	8	4221.4	98.5	4335	24	AAL49815	Human platelet all
	9	4219.8	98.4	4335	24	AAL49816	Human platelet all
	10	4219.6	98.4	4473	21	AAA62010	Hydrophobic domain
	11	4216.6	98.4	4335	21	AAA60199	Hydrophobic domain
	12	4140.2	96.6	4146	25	AAD49436	Human blood cell s
	13	4082.4	95.2	4197	25	AAD49437	Human r150 DNA #2.
	14	3187.4	74.4	3535	24	ABQ79968	Human CD109 K15 pr
	15	3185.8	74.3	3535	24	ABQ79969	Human CD109 K15 va
	16	1017	23.7	1448	24	ABL90478	Human polynucleoti
c	17	975.4	22.8	1459	22	ABA08825	Human secreted pro
	18	812.2	18.9	2403	24	ABV77330	Alpha-1 proteinase
	19	730.8	17.0	1300	22	AAF81735	Human protease and
	20	541.6	12.6	875	20	AAX20449	Human secreted pro
	21	525	12.2	821	25	ABX08843	Angiogenesis-assoc
	22	497.6	11.6	744	20	AAZ15521	Human gene express
	23	331.6	7.7	690	23	AAS70864	DNA encoding novel
	24	300	7.0	300	20	AAZ13539	Human gene express
	25	272	6.3	354	25	ABX47702	Bovine EST associa
	26	175	4.1	744	20	AAZ16725	Human gene express
	27	174	4.1	4527	22	AAC90951	C. elegans alpha-2
	28	174	4.1	4560	22	AAC90952	C. elegans alpha-2
	29	172.6	4.0	2285	23	ABL20611	Drosophila melanog
	30	171.6	4.0	2480	23	ABL12153	Drosophila melanog
	31	152.2	3.6	3612	23	ABL19705	Drosophila melanog
c	32	143.8	3.4	6324	23	ABL19704	Drosophila melanog
c	33	127.2	3.0	5611	23	ABL20610	Drosophila melanog
c	34	127.2	3.0	6979	23	ABL12152	Drosophila melanog
	35	126.8	3.0	4615	23	AAS92193	DNA encoding novel
	36	126.8	3.0	4615	24	ABN97232	Gene #3730 used to
	37	126.8	3.0	4629	23	AAS92192	DNA encoding novel
	38	126.8	3.0	5816	21	AAC74484	Human ORFX ORF39 p
	39	125.2	2.9	4487	24	ABK92035	DNA encoding novel
	40	125.2	2.9	4488	24	ABS59327	Human alpha-2-macr
	41	120.8	2.8	4660	24	ABQ93897	Human pregnancy zo
	42	120.6	2.8	2608	24	AAL49817	Human platelet all
	43	119.8	2.8	4501	25	ABX70464	DNA encoding human
	44	119.4	2.8	1140	22	ABA06696	Human cDNA SEQ ID
	45	119.4	2.8	1140	22	ABA06869	Human polynucleoti

ALIGNMENTS

```
RESULT 1
AAD49434
ID   AAD49434 standard; DNA; 4369 BP.
XX
AC   AAD49434;
XX
DT   24-MAR-2003  (first entry)
XX
DE   Human r150 gene #1.
XX
KW   Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW   glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW   therapy; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             26..4312
FT                   /*tag=  a
FT                   /product= "Human r1520 protein #1"
FT                   /transl_except= (pos:2132..2134, aa:Xaa)
FT                   /note= "Xaa corresponds to Ser, Tyr"
XX
PN   WO200285942-A2.
XX
PD   31-OCT-2002.
XX
PF   24-APR-2002; 2002WO-CA00560.
XX
PR   24-APR-2001; 2001US-285713P.
PR   14-FEB-2002; 2002US-356163P.
```

```
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   P-PSDB; AAE32012.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 91-93; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is a gene encoding
CC   human r150 protein.
XX
SQ   Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;

  Query Match            99.9%;  Score 4281.2;  DB 25;  Length 4369;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 4283;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 85

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 145

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 205

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 265

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 325

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 385

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 445

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 505

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 565

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 625

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATA[illegible]TATCAATCATTTCAG 660
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db        626 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAG 685

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 745

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTG 805

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 865

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 925

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 985

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1045

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1105

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1165

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1225

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1285

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1345

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1405

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1406 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1465

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1466 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1525

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1526 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1585

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1586 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1645

Qy       1621 GAAATTATAAGTGATGTTCTAAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1646 GAAATTATAAGTGATGTTCTAAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1705

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1706 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1765
```

Jan 20 08:47:21

```
Qy   1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1766 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1825

Qy   1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1826 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1885

Qy   1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1886 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1945

Qy   1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1946 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2005

Qy   1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2006 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2065

Qy   2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2066 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2125

Qy   2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2126 ATGGGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2185

Qy   2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2186 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2245

Qy   2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2246 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2305

Qy   2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2306 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2365

Qy   2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db   2366 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2425

Qy   2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2426 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2485

Qy   2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2486 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2545

Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2546 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2605

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2606 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2665

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2666 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2725

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2726 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2785

Qy   2761 TGTGGTGAACAGAACATGATAAATTTGCTCCAAATATTTACATTTGGATTATCTGACT 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2786 TGTGGTGAACAGAACATGATAAATTTGCTCCAAATATTTACATTTGGATTATCTGACT 2845

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2846 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2905

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2906 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2965

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2966 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3025

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3026 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3085

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3086 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3145

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3146 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3205

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3206 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3265

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3266 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3325

Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3326 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3385

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3386 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3445

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3446 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3505

Qy   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3506 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3565

Qy   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3566 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3625

Qy   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTG 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3626 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTG 3685

Qy   3661 GTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTC 3720
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3686 GTACAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTC 3745

Qy   3721 AATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAAT 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3746 AATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAAT 3805

Qy   3781 CAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTG 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3806 CAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTG 3865

Qy   3841 GATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAA 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3866 GATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAA 3925

Qy   3901 GTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTG 3960
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3926 GTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTG 3985

Qy         3961 AAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACC 4020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         3986 AAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACC 4045

Qy         4021 CAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCT 4080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4046 CAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCT 4105

Qy         4081 TCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCT 4140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4106 TCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCT 4165

Qy         4141 GAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAG 4200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4166 GAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAG 4225

Qy         4201 GATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTT 4260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         4226 GATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTT 4285

Qy         4261 CTGTACTTTATGGAACTTTGGCTGTGA 4287
                |||||||||||||||||||||||||||
Db         4286 CTGTACTTTATGGAACTTTGGCTGTGA 4312
```

```
RESULT 2
ABQ79964
ID   ABQ79964 standard; cDNA; 4761 BP.
XX
AC   ABQ79964;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1 protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82165.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 1a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1 cDNA sequence.
XX
SQ   Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

  Query Match             98.5%;  Score 4224.4;  DB 24;  Length 4761;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy            1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy           61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy          121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy          181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy          241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy          301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy          361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy          421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy          481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy          541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy          601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy          661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

GenCore version 5.1.6

nucleic search, using sw model

January 16, 2004, 12:38:20 ; Search time 14986 Seconds
(without alignments)
11702.894 Million cell updates/sec

cle: US-10-020-095-3
rfect score: 4287
Sequence: 1 atgcagggcccaccgctcct..........ttatggaactttggctgtga 4287

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	4224.4	98.5	4761	6	AX534934	AX534934 Sequence
	2	4224.4	98.5	5883	9	AF410459	AF410459 Homo sapi
	3	4224.4	98.5	5895	6	AX534938	AX534938 Sequence
	4	4222.8	98.5	4761	6	AX534936	AX534936 Sequence
	5	4222.8	98.5	5895	6	AX534940	AX534940 Sequence
	6	4221.4	98.5	4335	6	AX537610	AX537610 Sequence
	7	4219.8	98.4	4335	6	AX537612	AX537612 Sequence
	8	3187.4	74.4	3535	6	AX534942	AX534942 Sequence
	9	3185.8	74.3	3535	6	AX534944	AX534944 Sequence
	10	2709.2	63.2	4735	10	AY083458	AY083458 Mus muscu
	11	2709.2	63.2	5644	10	BC052443	BC052443 Mus muscu
	12	2448.2	57.1	2938	9	HSM805581	AL834478 Homo sapi
	13	1969.4	45.9	2273	9	AK095888	AK095888 Homo sapi
	14	730.8	17.0	1300	6	AX083146	AX083146 Sequence
	15	232	5.4	163577	9	AL590428	AL590428 Human DNA
	16	230	5.4	176144	2	AC012408	AC012408 Homo sapi
	17	219.6	5.1	5603	3	CIN431688	AJ431688 Ciona int
	18	203.8	4.8	4678	3	D83196	D83196 Limulus sp.
c	19	182	4.2	251250	2	AC097023	AC097023 Rattus no
	20	177.4	4.1	91419	9	AL591480	AL591480 Human DNA
	21	175	4.1	149597	2	AC034271	AC034271 Homo sapi
	22	174	4.1	4527	6	AX054953	AX054953 Sequence
	23	174	4.1	4560	6	AX054954	AX054954 Sequence
	24	171.6	4.0	2629	3	AY118302	AY118302 Drosophil
	25	167.6	3.9	4715	5	GGOVOM	X78801 G.gallus mR
	26	167	3.9	4578	3	AY122084	AY122084 Drosophil
	27	167	3.9	4579	3	DME269539	AJ269539 Drosophil
c	28	160	3.7	144589	2	AC112668	AC112668 Mus muscu
	29	160	3.7	191258	2	AC118254	AC118254 Mus muscu
	30	155.4	3.6	4620	10	RNA1I3	X52984 Rat mRNA fo
c	31	154.8	3.6	176144	2	AC012408	AC012408 Homo sapi
	32	154.2	3.6	4636	10	RATA1I3A	J03552 Rat plasma
	33	152.2	3.6	4220	3	DME269538	AJ269538 Drosophil
	34	151.2	3.5	4575	10	GPIMSPB	D84339 Cavia porce
	35	149.8	3.5	2501	10	RATA1I3	J03524 Rat alpha-1
	36	146.4	3.4	4950	5	XELENDO	L63543 Xenopus lae
c	37	143.8	3.4	122013	3	AC092399	AC092399 Drosophil
c	38	143.8	3.4	164414	3	AC092395	AC092395 Drosophil
c	39	143.8	3.4	168503	3	AC115484	AC115484 Drosophil
c	40	143.8	3.4	275699	3	AE003649	AE003649 Drosophil
c	41	143.8	3.4	337178	3	DROSADH08	AE003414 Drosophil
c	42	140.8	3.3	191258	2	AC118254	AC118254 Mus muscu
	43	136	3.2	4623	10	BC036299	BC036299 Mus muscu
	44	136	3.2	4672	10	BC051037	BC051037 Mus muscu
	45	134.8	3.1	2689	5	AB026130	AB026130 Cyprinus

ALIGNMENTS

RESULT 1
AX534934
LOCUS AX534934 4761 bp DNA linear PAT 22-NOV-2002
DEFINITION Sequence 1 from Patent WO02070696.
ACCESSION AX534934
VERSION AX534934.1 GI:25261474
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Schuh,A. and Sutherland,R.D.
TITLE Cd109 nucleic acid molecules polypeptides and methods of use
JOURNAL Patent: WO 02070696-A 1 12-SEP-2002;

```
Qy     3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy     3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy     3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy     3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy     3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy     3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy     4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy     4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy     4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy     4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy     4270 ATGGAACTTTGGCTGTGA 4287
            ||||||||||||||||||
Db     4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 2
AF410459
LOCUS       AF410459                5883 bp    mRNA    linear   PRI 02-MAR-2002
DEFINITION  Homo sapiens CD109 (CD109) mRNA, complete cds.
ACCESSION   AF410459
VERSION     AF410459.1  GI:19071208
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5883)
  AUTHORS   Lin,M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
            Wu,X.F. and Schuh,A.C.
  TITLE     Cell surface antigen CD109 is a novel member of the alpha(2)
            macroglobulin/C3, C4, C5 family of thioester-containing proteins
  JOURNAL   Blood 99 (5), 1683-1691 (2002)
  MEDLINE   21849742
   PUBMED   11861284
REFERENCE   2  (bases 1 to 5883)
  AUTHORS   Lin,M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
            Wu,X.-F. and Schuh,A.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's
            College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada
FEATURES             Location/Qualifiers
     source          1. .5883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="K1"
     gene            1. .5883
                     /gene="CD109"
     5'UTR           1. .112
                     /gene="CD109"
     CDS             113. .4450
                     /gene="CD109"
                     /note="associated with the Gov alloantigen system"
                     /codon_start=1
                     /product="CD109"
                     /protein_id="AAL84159.1"
                     /db_xref="GI:19071209"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
     3'UTR           4451. .5883
                     /gene="CD109"
     misc_feature    4748
                     /gene="CD109"
                     /note="alternative polyA site found on clone K1"
BASE COUNT     1697 a   1173 c   1241 g   1772 t
ORIGIN

  Query Match             98.5%;  Score 4224.4;  DB 9;  Length 5883;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;
```

```
Qy        1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy       61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy      121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy      181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy      241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy      301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
```

```
Db    413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy    361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
Db    473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy    421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
Db    533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy    481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
Db    593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy    541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
Db    653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy    601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
Db    713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy    661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
Db    773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy    721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
Db    833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 892

Qy    781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
Db    893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 952

Qy    841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
Db    953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy    901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
Db   1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy    961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
Db   1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy   1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
Db   1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy   1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
Db   1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy   1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
Db   1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy   1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
Db   1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy   1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
Db   1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy   1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
Db   1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy   1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
Db   1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy   1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
Db   1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy   1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
Db   1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy   1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
Db   1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy   1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
Db   1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy   1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
Db   1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy   1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
Db   1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy   1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
Db   1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy   1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
Db   1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy   1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
Db   2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy   1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
Db   2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy   2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
Db   2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy   2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
Db   2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy   2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
Db   2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy   2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
Db   2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy   2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
Db   2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy   2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
Db   2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy   2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
Db   2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy   2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
Db   2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632
```

```
Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
Db   2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
Db   2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
Db   2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
Db   2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
Db   2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
Db   2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
Db   2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
Db   3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
Db   3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
Db   3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
Db   3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
Db   3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
Db   3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
Db   3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
Db   3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
Db   3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

     3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
          GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

          GGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
          CACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
Db   3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy   3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
Db   3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy   3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
Db   3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy   3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
Db   3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy   3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
Db   3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy   3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
Db   4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy   3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
Db   4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy   3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
Db   4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy   4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
Db   4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy   4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
Db   4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy   4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
Db   4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy   4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
Db   4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy   4270 ATGGAACTTTGGCTGTGA 4287
Db   4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 3
AX534938
LOCUS       AX534938                5895 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 5 from Patent WO02070696.
ACCESSION   AX534938
VERSION     AX534938.1  GI:25261482
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 5 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1. .5895
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
```

Same as Result 1

```
Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 6
AX537610
LOCUS       AX537610                4335 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO02070738.
ACCESSION   AX537610
VERSION     AX537610.1  GI:25269531
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Ouwehand,W.
  TITLE     Diagnosis and treatment of blood disorders
  JOURNAL   Patent: WO 02070738-A 1 12-SEP-2002;
            Schuh, Andre (CA)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..>4335
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57496.1"
                     /db_xref="GI:25269532"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
COUNT       1296 a    837 c    922 g   1280 t

atch              98.5%;  Score 4221.4;  DB 6;  Length 4335;
l Similarity      98.8%;  Pred. No. 0;
  Conservative       0;  Mismatches    1;  Indels   51;  Gaps    1;
```

```
Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

Qy    1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy    1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy    1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy    1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy    1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy    1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy    1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy    1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy    1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy    1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy    1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy    1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy    1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy    1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy    1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy    2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy    2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy    2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy    2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy    2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy    2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
           ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db    2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400

Qy    2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy    2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy    2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy    2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy    2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy    2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy    2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTGGATTATCTGACT 2820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTGGATTATCTGACT 2820

Qy    2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy    2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy    2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy    3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy    3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy    3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy    3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy    3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
```

```
Qy       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy       3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                        ||||||||||||||||||
Db       3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy       3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy       3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy       3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy       3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy       3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy       3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy       4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy       4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

         4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         4261 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

         4270 ATGGAACTTTGGCTG 4284
              |||||||||||||||
            1 ATGGAACTTTGGCTG 4335
```

```
RESULT 7
AX537612
LOCUS       AX537612               4335 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 3 from Patent WO02070738.
ACCESSION   AX537612
VERSION     AX537612.1  GI:25269535
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Ouwehand,W.
  TITLE     Diagnosis and treatment of blood disorders
  JOURNAL   Patent: WO 02070738-A 3 12-SEP-2002;
            Schuh, Andre (CA).
FEATURES             Location/Qualifiers
     source          1. .4335
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1. .>4335
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57497.1"
                     /db_xref="GI:25269536"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1295 a    838 c    922 g   1280 t
ORIGIN
```

SAme AS Result 6

```
  Query Match             98.4%;  Score 4219.8;  DB 6;  Length 4335;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4282;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 12:36:15 ; Search time 1017 Seconds
(without alignments)
11379.047 Million cell updates/sec

Title: US-10-020-095-3
Perfect score: 4287
Sequence: 1 atgcagggcccaccgctcct..........ttatggaactttggctgtga 4287

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	4281.2	99.9	4369	25	AAD49434	Human r150 gene #1
	2	4224.4	98.5	4761	24	ABQ79964	Human CD109 K1 pro
	3	4224.4	98.5	5882	25	AAD49440	Human blood cell s
	4	4224.4	98.5	5895	24	ABQ79966	Human CD109 K1-H7
	5	4223.4	98.5	5883	25	AAD49435	Human r150 DNA #1.
	6	4222.8	98.5	4761	24	ABQ79965	Human CD109 K1 var
	7	4222.8	98.5	5895	24	ABQ79967	Human CD109 K1-H7
	8	4221.4	98.5	4335	24	AAL49815	Human platelet all
	9	4219.8	98.4	4335	24	AAL49816	Human platelet all
	10	4219.6	98.4	4473	21	AAA62010	Hydrophobic domain
	11	4216.6	98.4	4335	21	AAA60199	Hydrophobic domain
	12	4140.2	96.6	4146	25	AAD49436	Human blood cell s
	13	4082.4	95.2	4197	25	AAD49437	Human r150 DNA #2.
	14	3187.4	74.4	3535	24	ABQ79968	Human CD109 K15 pr
	15	3185.8	74.3	3535	24	ABQ79969	Human CD109 K15 va
	16	1017	23.7	1448	24	ABL90478	Human polynucleoti
c	17	975.4	22.8	1459	22	ABA08825	Human secreted pro
	18	812.2	18.9	2403	24	ABV77330	Alpha-1 proteinase
	19	730.8	17.0	1300	22	AAF81735	Human protease and
	20	541.6	12.6	875	20	AAX20449	Human secreted pro
	21	525	12.2	821	25	ABX08843	Angiogenesis-assoc
	22	497.6	11.6	744	20	AAZ15521	Human gene express
	23	331.6	7.7	690	23	AAS70864	DNA encoding novel
	24	300	7.0	300	20	AAZ13539	Human gene express
	25	272	6.3	354	25	ABX47702	Bovine EST associa
	26	175	4.1	744	20	AAZ16725	Human gene express
	27	174	4.1	4527	22	AAC90951	C. elegans alpha-2
	28	174	4.1	4560	22	AAC90952	C. elegans alpha-2
	29	172.6	4.0	2285	23	ABL20611	Drosophila melanog
	30	171.6	4.0	2480	23	ABL12153	Drosophila melanog
	31	152.2	3.6	3612	23	ABL19705	Drosophila melanog
c	32	143.8	3.4	6324	23	ABL19704	Drosophila melanog
c	33	127.2	3.0	5611	23	ABL20610	Drosophila melanog
c	34	127.2	3.0	6979	23	ABL12152	Drosophila melanog
	35	126.8	3.0	4615	23	AAS92193	DNA encoding novel
	36	126.8	3.0	4615	24	ABN97232	Gene #3730 used to
	37	126.8	3.0	4629	23	AAS92192	DNA encoding novel
	38	126.8	3.0	5816	21	AAC74484	Human ORFX ORF39 p
	39	125.2	2.9	4487	24	ABK92035	DNA encoding novel
	40	125.2	2.9	4488	24	ABS59327	Human alpha-2-macr
	41	120.8	2.8	4660	24	ABQ93897	Human pregnancy zo
	42	120.6	2.8	2608	24	AAL49817	Human platelet all
	43	119.8	2.8	4501	25	ABX70464	DNA encoding human
	44	119.4	2.8	1140	22	ABA06696	Human cDNA SEQ ID
	45	119.4	2.8	1140	22	ABA06869	Human polynucleoti

ALIGNMENTS

RESULT 1
AAD49434
~~ID~~ AAD49434 standard; DNA; 4369 BP.
XX
AC AAD49434;
XX
DT 24-MAR-2003 (first entry)
XX
DE Human r150 gene #1.
XX
KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW therapy; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 26..4312
FT /*tag= a
FT /product= "Human r1520 protein #1"
FT /transl_except= (pos:2132..2134, aa:Xaa)
FT /note= "Xaa corresponds to Ser, Tyr"
XX
PN WO200285942-A2.
XX
PD 31-OCT-2002.
XX
PF 24-APR-2002; 2002WO-CA00560.
XX
PR 24-APR-2001; 2001US-285713P.
PR 14-FEB-2002; 2002US-356163P.

```
XX
PA   (UYMC-) UNIV MCGILL.
XX
PI   Philip A,  Tam B;
XX
DR   WPI; 2003-093100/08.
DR   P-PSDB; AAE32012.
XX
PT   Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT   comprises r150 protein which acts as accessory receptor of TGF-beta,
PT   useful for negatively modulating TGF-beta activity, and thus for
PT   treating cancer  -
XX
PS   Claim 7; Page 91-93; 127pp; English.
XX
CC   The invention relates to novel transforming growth factor (TGF)-beta1
CC   binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-
CC   anchored TGF-beta1 binding protein referred to as r150 which acts as
CC   an accessory receptor of TGF-beta. The invention is used for negatively
CC   modulating TGF-beta activity, and thus for treating conditions
CC   characterised by overproduction of TGF-beta, such as cancer. Antisense
CC   molecules of the invention are used for increasing TGF-beta availability
CC   and increase graft success. The present sequence is a gene encoding
CC   human r150 protein.
XX
SQ   Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;

  Query Match            99.9%;  Score 4281.2;  DB 25;  Length 4369;
  Best Local Similarity  99.9%;  Pred. No. 0;
  Matches 4283;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         26 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 85

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 145

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 205

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 265

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 325

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 385

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 445

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 505

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 565

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        566 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 625

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATA[illegible]TATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        626 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATACTATCAATCATTTCAG 685

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        686 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 745

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        746 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 805

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        806 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 865

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        866 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 925

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        926 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 985

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        986 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1045

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1046 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1105

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1106 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1165

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1166 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1225

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1226 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1285

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1286 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1345

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1346 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1405

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1406 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1465

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1466 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1525

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1526 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1585

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1586 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1645

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1646 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1705

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1706 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1765
```

```
Qy   1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
Db   1766 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1825

Qy   1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
Db   1826 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1885

Qy   1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
Db   1886 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1945

Qy   1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
Db   1946 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2005

Qy   1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
Db   2006 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2065

Qy   2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
Db   2066 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2125

Qy   2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
Db   2126 ATGGGTTNCAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2185

Qy   2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
Db   2186 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2245

Qy   2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
Db   2246 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2305

Qy   2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
Db   2306 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2365

Qy   2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
Db   2366 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2425

Qy   2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
Db   2426 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2485

Qy   2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
Db   2486 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2545

Qy   2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
Db   2546 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2605

Qy   2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
Db   2606 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2665

Qy   2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
Db   2666 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2725

Qy   2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
Db   2726 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2785

Qy   2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
Db   2786 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2845

Qy   2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
Db   2846 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2905

Qy   2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
Db   2906 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2965

Qy   2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
Db   2966 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3025

Qy   3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
Db   3026 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3085

Qy   3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
Db   3086 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3145

Qy   3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
Db   3146 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3205

Qy   3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
Db   3206 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3265

Qy   3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
Db   3266 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3325

Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
Db   3326 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3385

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
Db   3386 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3445

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
Db   3446 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3505

Qy   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
Db   3506 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3565

Qy   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
Db   3566 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3625

Qy   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTG 3660
Db   3626 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTCTTGCTGTG 3685

Qy   3661 GTACAGCCAATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTC 3720
Db   3686 GTACAGCCAACGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTC 3745

Qy   3721 AATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAAT 3780
Db   3746 AATGTTGTATATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAAT 3805

Qy   3781 CAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTG 3840
Db   3806 CAAGAAGCCTTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTG 3865

Qy   3841 GATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAA 3900
Db   3866 GATTTGAATGTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAA 3925

Qy   3901 GTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTG 3960
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3926 GTTAACCTATTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTG 3985

Qy       3961 AAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACC 4020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3986 AAGAAAGTGGAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACC 4045

Qy       4021 CAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCT 4080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4046 CAGTTTTGTGTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCT 4105

Qy       4081 TCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCT 4140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4106 TCAGTGTCCATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCT 4165

Qy       4141 GAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAG 4200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4166 GAAGTGAAGCTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAG 4225

Qy       4201 GATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTT 4260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4226 GATGGAGCTTCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTCTGTTTCAAGCTT 4285

Qy       4261 CTGTACTTTATGGAACTTTGGCTGTGA 4287
              |||||||||||||||||||||||||||
Db       4286 CTGTACTTTATGGAACTTTGGCTGTGA 4312
```

```
RESULT 2
ABQ79964
ID   ABQ79964 standard; cDNA; 4761 BP.
XX
AC   ABQ79964;
XX
DT   23-DEC-2002  (first entry)
XX
DE   Human CD109 K1 protein encoding cDNA.
XX
KW   CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;
KW   immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW   cardiovascular; vasotropic; gene therapy; CD109 K1; gene; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             113..4450
FT                   /*tag=  a
FT                   /product= "CD109 K1"
XX
PN   WO200270696-A2.
XX
PD   12-SEP-2002.
XX
PF   07-MAR-2002; 2002WO-CA00292.
XX
PR   07-MAR-2001; 2001US-273814P.
XX
PA   (SCHU/) SCHUH A.
PA   (SUTH/) SUTHERLAND R D.
XX
PI   Schuh A,  Sutherland RD;
XX
DR   WPI; 2002-713450/77.
DR   P-PSDB; ABB82165.
XX
PT   New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT   particularly for treating strokes, myocardial infarctions, thrombosis,
PT   thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT   transplantation -
XX
PS   Claim 1; Fig 1a; 156pp; English.
XX
CC   The invention relates to isolated nucleic acid molecules encoding CD109
CC   polypeptides. These nucleic acid molecules include the human cDNA
CC   sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.
CC   CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC   family of thioester-containing proteins.The CD109 polypeptides can be
CC   expressed by standard recombinant methodology. The CD109 nucleic acid,
CC   CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC   for treating or detecting a disease or disorder, e.g. conditions
CC   associated with endothelial activation, platelet activation, activation
CC   of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC   and of the complement system, quantitative or qualitative abnormalities
CC   of platelet function, increased or impaired platelet aggregation and
CC   activation, increased or impaired activation of the coagulation and/or
CC   fibrinolytic systems, or impaired or increased immune activation. These
CC   are also useful for treating cardiovascular disorders, stroke, myocardial
CC   infarction, thrombosis, embolism, peripheral vascular disease,
CC   thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC   transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC   or their antisense nucleotide sequence are useful in gene therapy for
CC   treating these conditions. The present sequence represents the human
CC   CD109 K1 cDNA sequence.
XX
SQ   Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

  Query Match             98.5%;  Score 4224.4;  DB 24;  Length 4761;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

GenCore version 5.1.6

- nucleic search, using sw model

```
         January 16, 2004, 12:38:20 ; Search time 14986 Seconds
                                     (without alignments)
                                     11702.894 Million cell updates/sec

   e:                US-10-020-095-3
 fect score:         4287
 quence:             1 atgcagggcccaccgctcct..........ttatggaactttggctgtga 4287

 Scoring table:      IDENTITY_NUC
                     Gapop 10.0 , Gapext 1.0

 Searched:           2888711 seqs, 20454813386 residues

 Total number of hits satisfying chosen parameters:      5777422

 Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

 Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

 Database :      GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*
                 15:  em_ba:*
                 16:  em_fun:*
                 17:  em_hum:*
                 18:  em_in:*
                 19:  em_mu:*
                 20:  em_om:*
                 21:  em_or:*
                 22:  em_ov:*
                 23:  em_pat:*
                 24:  em_ph:*
                 25:  em_pl:*
                 26:  em_ro:*
                 27:  em_sts:*
                 28:  em_un:*
                 29:  em_vi:*
                 30:  em_htg_hum:*
                 31:  em_htg_inv:*
                 32:  em_htg_other:*
                 33:  em_htg_mus:*
                 34:  em_htg_pln:*
                 35:  em_htg_rod:*
                 36:  em_htg_mam:*
                 37:  em_htg_vrt:*
                 38:  em_sy:*
                 39:  em_htgo_hum:*
                 40:  em_htgo_mus:*
                 41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	4224.4	98.5	4761	6	AX534934	AX534934 Sequence
	2	4224.4	98.5	5883	9	AF410459	AF410459 Homo sapi
	3	4224.4	98.5	5895	6	AX534938	AX534938 Sequence
	4	4222.8	98.5	4761	6	AX534936	AX534936 Sequence
	5	4222.8	98.5	5895	6	AX534940	AX534940 Sequence
	6	4221.4	98.5	4335	6	AX537610	AX537610 Sequence
	7	4219.8	98.4	4335	6	AX537612	AX537612 Sequence
	8	3187.4	74.4	3535	6	AX534942	AX534942 Sequence
	9	3185.8	74.3	3535	6	AX534944	AX534944 Sequence
	10	2709.2	63.2	4735	10	AY083458	AY083458 Mus muscu
	11	2709.2	63.2	5644	10	BC052443	BC052443 Mus muscu
	12	2448.2	57.1	2938	9	HSM805581	AL834478 Homo sapi
	13	1969.4	45.9	2273	9	AK095888	AK095888 Homo sapi
	14	730.8	17.0	1300	6	AX083146	AX083146 Sequence
	15	232	5.4	163577	9	AL590428	AL590428 Human DNA
	16	230	5.4	176144	2	AC012408	AC012408 Homo sapi
	17	219.6	5.1	5603	3	CIN431688	AJ431688 Ciona int
	18	203.8	4.8	4678	3	D83196	D83196 Limulus sp.
c	19	182	4.2	251250	2	AC097023	AC097023 Rattus no
	20	177.4	4.1	91419	9	AL591480	AL591480 Human DNA
	21	175	4.1	149597	2	AC034271	AC034271 Homo sapi
	22	174	4.1	4527	6	AX054953	AX054953 Sequence
	23	174	4.1	4560	6	AX054954	AX054954 Sequence
	24	171.6	4.0	2629	3	AY118302	AY118302 Drosophil
	25	167.6	3.9	4715	5	GGOVOM	X78801 G.gallus mR
	26	167	3.9	4578	3	AY122084	AY122084 Drosophil
	27	167	3.9	4579	3	DME269539	AJ269539 Drosophil
c	28	160	3.7	144589	2	AC112668	AC112668 Mus muscu
	29	160	3.7	191258	2	AC118254	AC118254 Mus muscu
	30	155.4	3.6	4620	10	RNA1I3	X52984 Rat mRNA fo
c	31	154.8	3.6	176144	2	AC012408	AC012408 Homo sapi
	32	154.2	3.6	4636	10	RATA1I3A	J03552 Rat plasma
	33	152.2	3.6	4220	3	DME269538	AJ269538 Drosophil
	34	151.2	3.5	4575	10	GPIMSPB	D84339 Cavia porce
	35	149.8	3.5	2501	10	RATA1I3	J03524 Rat alpha-1
	36	146.4	3.4	4950	5	XELENDO	L63543 Xenopus lae
c	37	143.8	3.4	122013	3	AC092399	AC092399 Drosophil
c	38	143.8	3.4	164414	3	AC092395	AC092395 Drosophil
c	39	143.8	3.4	168503	3	AC115484	AC115484 Drosophil
c	40	143.8	3.4	275699	3	AE003649	AE003649 Drosophil
c	41	143.8	3.4	337178	3	DROSADH08	AE003414 Drosophil
c	42	140.8	3.3	191258	2	AC118254	AC118254 Mus muscu
	43	136	3.2	4623	10	BC036299	BC036299 Mus muscu
	44	136	3.2	4672	10	BC051037	BC051037 Mus muscu
	45	134.8	3.1	2689	5	AB026130	AB026130 Cyprinus

ALIGNMENTS

```
RESULT 1
AX534934
LOCUS       AX534934                 4761 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 1 from Patent WO02070696.
ACCESSION   AX534934
VERSION     AX534934.1  GI:25261474
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 1 12-SEP-2002;
```

```
Qy     3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy     3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy     3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy     3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy     3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy     3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy     4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy     4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy     4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy     4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy     4270 ATGGAACTTTGGCTGTGA 4287
            ||||||||||||||||||
Db     4433 ATGGAACTTTGGCTGTGA 4450

RESULT 2
AF410459
LOCUS       AF410459                5883 bp    mRNA    linear   PRI 02-MAR-2002
DEFINITION  Homo sapiens CD109 (CD109) mRNA, complete cds.
ACCESSION   AF410459
VERSION     AF410459.1  GI:19071208
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5883)
  AUTHORS   Lin,M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
            Wu,X.F. and Schuh,A.C.
  TITLE     Cell surface antigen CD109 is a novel member of the alpha(2)
            macroglobulin/C3, C4, C5 family of thioester-containing proteins
  JOURNAL   Blood 99 (5), 1683-1691 (2002)
  MEDLINE   21849742
   PUBMED   11861284
REFERENCE   2  (bases 1 to 5883)
  AUTHORS   Lin,M., Sutherland,D.R., Horsfall,W., Totty,N., Yeo,E., Nayar,R.,
            Wu,X.-F. and Schuh,A.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's
            College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada
FEATURES             Location/Qualifiers
     source          1. .5883
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="K1"
     gene            1. .5883
                     /gene="CD109"
     5'UTR           1. .112
                     /gene="CD109"
     CDS             113. .4450
                     /gene="CD109"
                     /note="associated with the Gov alloantigen system"
                     /codon_start=1
                     /product="CD109"
                     /protein_id="AAL84159.1"
                     /db_xref="GI:19071209"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
     3'UTR           4451. .5883
                     /gene="CD109"
     misc_feature    4748
                     /gene="CD109"
                     /note="alternative polyA site found on clone K1"
BASE COUNT     1697 a   1173 c   1241 g   1772 t
ORIGIN

  Query Match            98.5%;  Score 4224.4;  DB 9;  Length 5883;
  Best Local Similarity  98.8%;  Pred. No. 0;
  Matches 4286;  Conservative    0;  Mismatches    1;  Indels   51;  Gaps    1;

Qy        1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      113 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 172

Qy       61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      173 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 232

Qy      121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      233 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 292

Qy      181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      293 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 352

Qy      241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 412

Qy      301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        413 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 472

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        473 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 532

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        533 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        593 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 652

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        653 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 712

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        713 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 772

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        773 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 832

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        833 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGAAGCCAGTG 892

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        893 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAAATATT 952

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        953 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 1012

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1013 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 1072

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1073 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1132

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1133 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1192

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1193 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1252

Qy       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1253 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1312

Qy       1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1313 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1372

Qy       1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1373 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1432

Qy       1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1433 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1492

Qy       1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1493 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1552

Qy       1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1553 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1612

Qy       1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1613 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1672

Qy       1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1673 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1732

Qy       1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1733 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1792

Qy       1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1793 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1852

Qy       1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1853 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1912

Qy       1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1913 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1972

Qy       1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1973 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 2032

Qy       1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2033 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 2092

Qy       1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2093 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2152

Qy       2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2153 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212

Qy       2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2213 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2272

Qy       2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2273 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2332

Qy       2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2333 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2392

Qy       2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2393 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2452

Qy       2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db       2453 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2512

Qy       2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2513 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2572

Qy       2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2573 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2632
```

```
Qy     2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2633 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692

Qy     2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2693 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2752

Qy     2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2753 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2812

Qy     2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2813 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2872

Qy     2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2873 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932

Qy     2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2933 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2992

Qy     2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2993 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 3052

Qy     2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3053 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3112

Qy     3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3113 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3172

Qy     3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3173 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3232

Qy     3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3233 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3292

Qy     3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3293 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3352

Qy     3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3353 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3412

Qy     3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3413 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3472

Qy     3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3473 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3532

Qy     3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3533 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3592

Qy     3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          GGAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3652

Qy          ACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
            ||||||||||||||||||||||||||||||||||||||||||||||||
Db          AAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3712

Qy     3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3713 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3772

Qy     3652 ------------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                      ||||||||||||||||||
Db     3773 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3832

Qy     3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3833 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892

Qy     3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3893 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3952

Qy     3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3953 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 4012

Qy     3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4013 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 4072

Qy     3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4073 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4132

Qy     3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4133 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4192

Qy     4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4252

Qy     4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4253 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4312

Qy     4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy     4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy     4270 ATGGAACTTTGGCTGTGA 4287
            ||||||||||||||||||
Db     4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 3
AX534938
LOCUS       AX534938                 5895 bp    DNA     linear   PAT 22-NOV-2002
DEFINITION  Sequence 5 from Patent WO02070696.
ACCESSION   AX534938
VERSION     AX534938.1  GI:25261482
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Sutherland,R.D.
  TITLE     Cd109 nucleic acid molecules polypeptides and methods of use
  JOURNAL   Patent: WO 02070696-A 5 12-SEP-2002;
            Schuh, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES             Location/Qualifiers
     source          1..5895
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
```

Same as Result 1

```
Qy       4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4313 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4372

Qy       4210 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4373 TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432

Qy       4270 ATGGAACTTTGGCTGTGA 4287
              ||||||||||||||||||
Db       4433 ATGGAACTTTGGCTGTGA 4450
```

```
RESULT 6
AX537610
LOCUS       AX537610                4335 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 1 from Patent WO02070738.
ACCESSION   AX537610
VERSION     AX537610.1  GI:25269531
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Ouwehand,W.
  TITLE     Diagnosis and treatment of blood disorders
  JOURNAL   Patent: WO 02070738-A 1 12-SEP-2002;
            Schuh, Andre (CA)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..>4335
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57496.1"
                     /db_xref="GI:25269532"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGYRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
     1296 a    837 c    922 g    1280 t

                 98.5%;  Score 4221.4;  DB 6;  Length 4335;
   rity          98.8%;  Pred. No. 0;
   rvative       0;  Mismatches    1;  Indels   51;  Gaps    1;
```

```
Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240

Qy        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACCTCTGAACAGTGCAGATGAG 300

Qy        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ATTTATGAGCTACGTGTAACCGGACGTACCCAGGATGAGATTTTATTCTCTAATAGTACC 360

Qy        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGCTTATCATTTGAGACCAAGAGAATATCTGTCTTCATTCAAACAGACAAGGCCTTATAC 420

Qy        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AAGCCAAAGCAAGAAGTGAAGTTTCGCATTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480

Qy        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAAACCTCTTTAAACATTCTCATTAAGGACCCCAAATCAAATTTGATCCAACAGTGGTTG 540

Qy        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TCACAACAAAGTGATCTTGGAGTCATTTCCAAAACTTTTCAGCTATCTTCCCATCCAATA 600

Qy        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTGGTGACTGGTCTATTCAAGTTCAAGTGAATGACCAGACATATTATCAATCATTTCAG 660

Qy        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GTTTCAGAATATGTATTACCAAAATTTGAAGTGACTTTGCAGACACCATTATATTGTTCT 720

Qy        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATGAATTCTAAGCATTTAAATGGTACCATCACGGCAAAGTATACATATGGGAAGCCAGTG 780

Qy        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAAGGAGACGTAACGCTTACATTTTTACCTTTATCCTTTTGGGGAAAGAAGAAAAATATT 840

Qy        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 ACAAAAACATTTAAGATAAATGGATCTGCAAACTTCTCTTTTAATGATGAAGAGATGAAA 900

Qy        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AATGTAATGGATTCTTCAAATGGACTTTCTGAATACCTGGATCTATCTTCCCCTGGACCA 960

Qy        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GTAGAAATTTTAACCACAGTGACAGAATCAGTTACAGGTATTTCAAGAAATGTAAGCACT 1020

Qy       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATGTGTTCTTCAAGCAACATGATTACATCATTGAGTTTTTTGATTATACTACTGTCTTG 1080

Qy       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
       1081 AAGCCATCTCTCAACTTCACAGCCACTGTGAAGGTAACTCGTGCTGATGGCAACCAACTG 1140

       1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 ACTCTTGAAGAAAGAAGAAATAATGTAGTCATAACAGTGACACAGAGAAACTATACTGAG 1200

Qy     1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TACTGGAGCGGATCTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAGAAAATAAATTAT 1260

Qy     1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 ACTGTCCCCCAAAGTGGAACTTTTAAGATTGAATTCCCAATCCTGGAGGATTCCAGTGAG 1320

Qy     1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 CTACAGTTGAAGGCCTATTTCCTTGGTAGTAAAAGTAGCATGGCAGTTCATAGTCTGTTT 1380

Qy     1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 AAGTCTCCTAGTAAGACATACATCCAACTAAAAACAAGAGATGAAAATATAAAGGTGGGA 1440

Qy     1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 TCGCCTTTTGAGTTGGTGGTTAGTGGCAACAAACGATTGAAGGAGTTAAGCTATATGGTA 1500

Qy     1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAACAAAATTCAACAATGTTCTCTTTAACA 1560

Qy     1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 CCAGAAAATTCTTGGACTCCAAAAGCCTGTGTAATTGTGTATTATATTGAAGATGATGGG 1620

Qy     1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 GAAATTATAAGTGATGTTCTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATAAGATAAAG 1680

Qy     1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 CTATATTGGAGTAAAGTGAAAGCTGAACCATCTGAGAAAGTCTCTCTTAGGATCTCTGTG 1740

Qy     1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 ACACAGCCTGACTCCATAGTTGGGATTGTAGCTGTTGACAAAAGTGTGAATCTGATGAAT 1800

Qy     1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 GCCTCTAATGATATTACAATGGAAAATGTGGTCCATGAGTTGGAACTTTATAACACAGGA 1860

Qy     1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 TATTATTTAGGCATGTTCATGAATTCTTTTGCAGTCTTTCAGGAATGTGGACTCTGGGTA 1920

Qy     1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 TTGACAGATGCAAACCTCACGAAGGATTATATTGATGGTGTTTATGACAATGCAGAATAT 1980

Qy     1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 GCTGAGAGGTTTATGGAGGAAAATGAAGGACATATTGTAGATATTCATGACTTTTCTTTG 2040

Qy     2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 GGTAGCAGTCCACATGTCCGAAAGCATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100

Qy     2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 ATGGGTTACAGGATTTACCAAGAATTTGAAGTAACTGTACCTGATTCTATCACTTCTTGG 2160

Qy     2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 GTGGCTACTGGTTTTGTGATCTCTGAGGACCTGGGTCTTGGACTAACAACTACTCCAGTG 2220

Qy     2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2221 GAGCTCCAAGCCTTCCAACCATTTTTCATTTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280

Qy     2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2281 GGTGAAGAATTTGCTTTGGAAATAACTATATTCAATTATTTGAAAGATGCCACTGAGGTT 2340

Qy     2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAGTGAAATAAAT 2400
            ||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     2341 AAGGTAATCATTGAGAAAAGTGACAAATTTGATATTCTAATGACTTCAAATGAAATAAAT 2400

Qy     2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2401 GCCACAGGCCACCAGCAGACCCTTCTGGTTCCCAGTGAGGATGGGGCAACTGTTCTTTTT 2460

Qy     2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2461 CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCACAGTCACAGCTCTTTCACCCACT 2520

Qy     2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2521 GCTTCTGATGCTGTCACCCAGATGATTTTAGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580

Qy     2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2581 TCACAATCCATCTTATTAGACTTGACTGACAATAGGCTACAGAGTACCCTGAAAACTTTG 2640

Qy     2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2641 AGTTTCTCATTTCCTCCTAATACAGTGACTGGCAGTGAAAGAGTTCAGATCACTGCAATT 2700

Qy     2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2701 GGAGATGTTCTTGGTCCTTCCATCAATGGCTTAGCCTCATTGATTCGGATGCCTTATGGC 2760

Qy     2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2761 TGTGGTGAACAGAACATGATAAATTTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820

Qy     2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2821 AAAAAGAAACAACTGACAGATAATTTGAAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880

Qy     2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2881 TACCAGAGAGAACTTCTCTATCAGAGGGAAGATGGCTCTTTCAGTGCTTTTGGGAATTAT 2940

Qy     2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2941 GACCCTTCTGGGAGCACTTGGTTGTCAGCTTTTGTTTTAAGATGTTTCCTTGAAGCCGAT 3000

Qy     3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3001 CCTTACATAGATATTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGGACAT 3060

Qy     3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3061 CAGAAATCCAACGGTGAATTTTGGGATCCAGGAAGAGTGATTCATAGTGAGCTTCAAGGT 3120

Qy     3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3121 GGCAATAAAAGTCCAGTAACACTTACAGCCTATATTGTAACTTCTCTCCTGGGATATAGA 3180

Qy     3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3181 AAGTATCAGCCTAACATTGATGTGCAAGAGTCTATCCATTTTTTGGAGTCTGAATTCAGT 3240

Qy     3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3241 AGAGGAATTTCAGACAATTATACTCTAGCCCTTATAACTTATGCATTGTCATCAGTGGGG 3300
```

```
Qy   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 AGTCCTAAAGCGAAGGAAGCTTTGAATATGCTGACTTGGAGAGCAGAACAAGAAGGTGGC 3360

Qy   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 ATGCAATTCTGGGTGTCATCAGAGTCCAAACTTTCTGACTCCTGGCAGCCACGCTCCCTG 3420

Qy   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAGACTTCTGAG 3480

Qy   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 GGAATCCCAATTATGAGGTGGCTAAGCAGGCAAAGAAATAGCTTGGGTGGTTTTGCATCT 3540

Qy   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAGCCCTAATGAATACA 3600

Qy   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCT--------- 3651
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 GAAAGGACAAATATCCAAGTGACCGTGACGGGGCCTAGCTCACCAAGTCCTGTAAAGTTT 3660

Qy   3652 -----------------------------------------CTTGCTGTGGTACAGCCA 3669
                                                   ||||||||||||||||||
Db   3661 CTGATTGACACACACAACCGCTTACTCCTTCAGACAGCAGAGCTTGCTGTGGTACAGCCA 3720

Qy   3670 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 ATGGCAGTTAATATTTCCGCAAATGGTTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3780

Qy   3730 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3789
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 TATAATGTGAAGGCTTCTGGGTCTTCTAGAAGACGAAGATCTATCCAAAATCAAGAAGCC 3840

Qy   3790 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3849
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 TTTGATTTAGATGTTGCTGTAAAAGAAAATAAAGATGATCTCAATCATGTGGATTTGAAT 3900

Qy   3850 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3909
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 GTGTGTACAAGCTTTTCGGGCCCGGGTAGGAGTGGCATGGCTCTTATGGAAGTTAACCTA 3960

Qy   3910 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 3969
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 TTAAGTGGCTTTATGGTGCCTTCAGAAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTG 4020

Qy   3970 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4029
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 GAATATGATCATGGAAAACTCAACCTCTATTTAGATTCTGTAAATGAAACCCAGTTTTGT 4080

Qy   4030 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4089
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAATACCCAAGATGCTTCAGTGTCC 4140

Qy   4090 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4149
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 ATAGTGGATTACTATGAGCCAAGGAGACAGGCGGTGAGAAGTTACAACTCTGAAGTGAAG 4200

Qy   4150 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4209
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 CTGTCCTCCTGTGACCTTTGCAGTGATGTCCAGGGCTGCCGTCCTTGTGAGGATGGAGCT 4260

Qy        TCAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        GAGGCTCCCATCATCACTCTTCAGTCATTTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4320

Qy        AACTTTGGCTG 4284
          |||||||||
Db        TTGGCTG 4335
```

RESULT 7
AX537612

```
LOCUS       AX537612                4335 bp    DNA     linear   PAT 23-NOV-2002
DEFINITION  Sequence 3 from Patent WO02070738.
ACCESSION   AX537612
VERSION     AX537612.1  GI:25269535
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Schuh,A. and Ouwehand,W.
  TITLE     Diagnosis and treatment of blood disorders
  JOURNAL   Patent: WO 02070738-A 3 12-SEP-2002;
            Schuh, Andre (CA)
FEATURES             Location/Qualifiers
     source          1..4335
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
     CDS             1..>4335
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD57497.1"
                     /db_xref="GI:25269536"
                     /translation="MQGPPLLTAAHLLCVCTAALAVAPGPRFLVTAPGIIRPGGNVTI
                     GVELLEHCPSQVTVKAELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIY
                     ELRVTGRTQDEILFSNSTRLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDFKPY
                     KTSLNILIKDPKSNLIQQWLSQQSDLGVISKTFQLSSHPILGDWSIQVQVNDQTYYQS
                     FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTITAKYTYGKPVKGDVTLTFLPLSFWGK
                     KKNITKTFKINGSANFSFNDEEMKNVMDSSNGLSEYLDLSSPGPVEILTTVTESVTGI
                     SRNVSTNVFFKQHDYIIEFFDYTTVLKPSLNFTATVKVTRADGNQLTLEERRNNVVIT
                     VTQRNYTEYWSGSNSGNQKMEAVQKINYTVPQSGTFKIEFPILEDSSELQLKAYFLGS
                     KSSMAVHSLFKSPSKTYIQLKTRDENIKVGSPFELVVSGNKRLKELSYMVVSRGQLVA
                     VGKQNSTMFSLTPENSWTPKACVIVYYIEDDGEIISDVLKIPVQLVFKNKIKLYWSKV
                     KAEPSEKVSLRISVTQPDSIVGIVAVDKSVNLMNASNDITMENVVHELELYNTGYYLG
                     MFMNSFAVFQECGLWVLTDANLTKDYIDGVYDNAEYAERFMEENEGHIVDIHDFSLGS
                     SPHVRKHFPETWIWLDTNMGSRIYQEFEVTVPDSITSWVATGFVISEDLGLGLTTTPV
                     ELQAFQPFFIFLNLPYSVIRGEEFALEITIFNYLKDATEVKVIIEKSDKFDILMTSNE
                     INATGHQQTLLVPSEDGATVLFPIRPTHLGEIPITVTALSPTASDAVTQMILVKAEGI
                     EKSYSQSILLDLTDNRLQSTLKTLSFSFPPNTVTGSERVQITAIGDVLGPSINGLASL
                     IRMPYGCGEQNMINFAPNIYILDYLTKKKQLTDNLKEKALSFMRQGYQRELLYQREDG
                     SFSAFGNYDPSGSTWLSAFVLRCFLEADPYIDIDQNVLHRTYTWLKGHQKSNGEFWDP
                     GRVIHSELQGGNKSPVTLTAYIVTSLLGYRKYQPNIDVQESIHFLESEFSRGISDNYT
                     LALITYALSSVGSPKAKEALNMLTWRAEQEGGMQFWVSSESKLSDSWQPRSLDIEVAA
                     YALLSHFLQFQTSEGIPIMRWLSRQRNSLGGFASTQDTTVALKALSEFAALMNTERTN
                     IQVTVTGPSSPSPVKFLIDTHNRLLLQTAELAVVQPMAVNISANGFGFAICQLNVVYN
                     VKASGSSRRRRSIQNQEAFDLDVAVKENKDDLNHVDLNVCTSFSGPGRSGMALMEVNL
                     LSGFMVPSEAISLSETVKKVEYDHGKLNLYLDSVNETQFCVNIPAVRNFKVSNTQDAS
                     VSIVDYYEPRRQAVRSYNSEVKLSSCDLCSDVQGCRPCEDGASGSHHHSSVIFIFCFK
                     LLYFMELWL"
BASE COUNT     1295 a    838 c    922 g   1280 t
ORIGIN
```

SAe AS Result 6

```
  Query Match             98.4%;  Score 4219.8;  DB 6;  Length 4335;
  Best Local Similarity   98.8%;  Pred. No. 0;
  Matches 4282;  Conservative    0;  Mismatches    2;  Indels   51;  Gaps    1;

Qy          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGCAGGGCCCACCGCTCCTGACCGCCGCCCACCTCCTCTGCGTGTGCACCGCCGCGCTG 60

Qy         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCCGTGGCTCCCGGGCCTCGGTTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 120

Qy        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AATGTGACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGACTGTGAAGGCG 180

Qy        181 GAGCTGCTCAAGACAGCATCAAACCTCACTGTCTCTGTCCTGGAAGCAGAAGGAGTCTTT 240
```

US 099155150FP1

Creation date: 02-19-2004
Indexing Officer: AGOMEZ - ALFREDO GOMEZ, JR.
Team: OIPEBackFileIndexing
Dossier: 09915515

Legal Date: 01-05-2004

No.	Doccode	Number of pages
1	SRNT	64

Total number of pages: 64

Remarks:

Order of re-scan issued on ....